# Patterns of habitat use by an insular endemic bird, the Azores woodpigeon (*Columba palumbus azorica*)

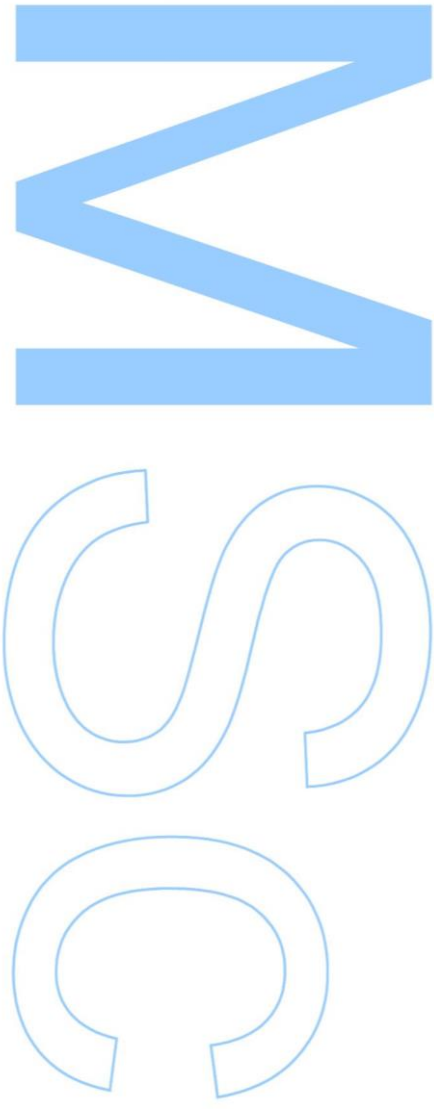

Júlio Daniel Gomes dos Santos
Mestrado em Biodiversidade, Genética e Evolução
Departamento de Biologia
2017

**Orientador**
José Carlos Brito, Faculty of Sciences and CIBIO-InBIO, University of Porto

**Coorientadores**
David Gonçalves, Faculty of Sciences and CIBIO-InBIO, University of Porto
Tiago M. Rodrigues, CIBIO-InBIO, University of Porto


Todas as correções determinadas pelo júri, e só essas, foram efetuadas.
O Presidente do Júri,

Porto, _____/_____/_____



# Acknowledgments

I would like to thank Rémi Fontaine, for all the hard work in collecting the data used in this study and to *Direção Regional da Ciência e Tecnologia* (Azores), for the financial support to the project "Ecology and evolution of the Azores woodpigeon: support tools for a management and conservation strategy" (AZORPI) (M2.1.2/I/025/2011) that was at the origin of the present study. I also must thank Maria Amélia Fonseca and Verónica Neves, researchers from the University of Azores, respectively principal investigator and team member of the AZORPI project.

I thank to CMMG (*Centro do Clima, Meteorologia e Mudanças Globais da Universidade dos Açores*), namely Professor Eduardo Brito de Azevedo, João Filipe Fernandes and Francisco Reis, for providing the environmental data used in this work.

I am very grateful to Professor José Alberto Gonçalves for his help on coordinate system conversion, to Cândida Vale for her guidance during my first steps modelling abundance distribution, and to all my master's degree colleagues that helped me during the last 2 years.

# Resumo

Cerca de 17% das espécies de aves existentes no mundo estão presentes em ilhas, e muitas são endémicas. No arquipélago dos Açores existe uma subespécie endémica do pombo-torcaz (*Columba palumbus*), o pombo-torcaz-dos-Açores (*C. p. azorica*). Este reproduz-se em todas as ilhas do arquipélago, mas é mais comum no Pico, Graciosa e Terceira. Apesar de ser relativamente comum no arquipélago, pouco se sabe sobre a sua biologia.

Os objetivos deste trabalho são investigar os padrões de distribuição da abundância e uso do habitat do pombo-torcaz-dos-Açores na ilha Terceira. Em particular pretendeu-se: i) investigar variações sazonais e diárias no uso do habitat com o intuito de perceber a dinâmica de movimentos da espécie; ii) mapear a distribuição da abundância e identificar fatores ambientais que podem estar relacionados com a distribuição da abundância observada do pombo-torcaz-dos-Açores; iii) modelar a distribuição potencial do pombo-torcaz-dos-Açores e comparar os resultados de três medidas de abundância diferentes.

A abundância do pombo-torcaz-dos-Açores (número de pombos / 10 minutos de observação) foi amostrada na ilha Terceira entre novembro de 2012 e agosto de 2014, em 100 pontos de amostragem, visitados quinzenalmente durante manhãs e tardes. Três índices de abundância foram gerados: *All data* (todos os dados), *Truncated data* (excluindo observações nulas) e *Maximum data* (abundância máxima registada). As diferenças na distribuição da abundância do pombo-torcaz-dos-Açores foram investigadas de acordo com variações diárias (Manhã vs. Tarde), sazonais (época de Reprodução vs. época de Não-reprodução) e anuais (Ano 1 vs. Ano 2). Apenas foram identificadas diferenças sazonais. Dada a correlação existente entre as variáveis ambientais (habitat, topografia e clima), foi usado o programa ArcGIS 10.1 para realizar uma Análise de Componentes Principais (PCA). Usando o R 3.3.2, determinaram-se regressões múltiplas para cada época (Reprodução e Não-reprodução) com o objetivo de avaliar relações entre as abundâncias observadas e os componentes da PCA.

Os dados *Truncated*, foram os que demonstraram uma relação mais forte com os três componentes da PCA (PC01, PC02 e PC03) incluídos nos modelos. O PC01 apresenta uma correlação negativa para o Índice de rugosidade do terreno (TRI) e a precipitação acumulada, o PC02 apresenta uma correlação positiva para as percentagens de florestas de coníferas e o PC03 apresenta uma correlação positiva

para a percentagem de vegetação natural e negativa para a percentagem de plantações de eucalipto. A abundância observada varia positivamente com a percentagem de florestas de coníferas e com a de plantações de eucaliptos, enquanto que a relação com o TRI, a precipitação acumulada e a vegetação natural é inversa. As regiões com valores de abundância previstos mais elevados correspondem a campos agrícolas, enquanto que as áreas com vegetação natural e áreas sociais e urbanas exibiram baixos valores de abundância esperada.

Os modelos desenvolvidos poderiam, eventualmente, ser melhorados se outras variáveis, como a distribuição e disponibilidade espacial de alimento, ou o índice da *Normalized Difference Vegetation Index* (NDVI), tivessem sido incluídas. Este é um dos primeiros estudos desta natureza para esta espécie, não apenas nos Açores mas também na Europa e muito trabalho tem que ser feito neste contexto.

**Palavras-chave**

*Columba palumbus*; Açores; modelos de regressão multipla; modelos de abundância; seleção de habitat

# Abstract

About 17% of the world's bird species are present on islands, and many are local endemics. The Azores archipelago is home to an endemic subspecies of woodpigeon (*Columba palumbus*), the Azores woodpigeon (*C. p. azorica*). The Azores woodpigeon breeds on all the islands of the archipelago, but is more common on Pico, Graciosa and Terceira islands. Despite being relatively common in the archipelago, little is known about the biology of this endemic bird.

The objectives of this work were to investigate patterns in the distribution of abundance and habitat use of the Azores woodpigeon on Terceira island. In particular, it was aimed to: i) investigate seasonal and daily variations in habitat use to understand the movement dynamics of the species; ii) map the distribution of abundance and identify environmental factors that may be related with observed distribution of the Azores woodpigeon abundance; iii) model the potential distribution of the Azores woodpigeon abundance and compare the results of three different measures of abundance.

Abundance data of the Azores woodpigeon (number of pigeons / 10 minutes of observation) was collected on Terceira island between November 2012 to August 2014, in 100 sampling points, visited fortnightly during mornings and afternoons. Three abundance datasets were generated: *All data*, *Truncated data* (excluding zero observations) and *Maximum data* (maximum recorded abundance). Differences in the distribution of abundance of the Azores woodpigeon were investigated according to daily variations (Morning vs. Afternoon), seasonal variation (Breeding vs. Non-breeding) and yearly variations (Year 1 vs. Year 2). Significant differences were only found between Breeding and Non-breeding seasons. Given the observed correlation among environmental variables (habitat, topography and climate), ArcGIS 10.1 was used to perform a principal component analysis (PCA). Using R 3.3.2 we performed multiple regressions models for the Breeding season and Non-breeding season to assess relationships between the observed abundances and the components of the PCA.

The *Truncated dataset* was the data transformation that exhibited the strongest relationship with the three PCA components (PC01, PC02 and PC03) retained in the models. PC01 shows a negative correlation in Terrain ruggedness index (TRI) and accumulated rainfall, PC02 shows a positive correlation in percentage of coniferous forests and PC03 shows a positive correlation in percentage of natural vegetation and a negative correlation in percentage of eucalyptus plantations. The observed abundance varies positively with the percentages of coniferous forests and eucalyptus

plantations, whereas the relation with TRI, accumulated rainfall and natural vegetation is inverse. The regions predicted with high values of abundance corresponded to agriculture fields, whereas areas with natural vegetation and social and urban areas exhibited lower values of expected abundance.

The models developed could be further improved if other variables, such as food distribution across space and normalized difference vegetation index, were included. This is one of the first studies of this nature for this species, not only in Azores but also in Europe and much more work need to be done.

## Keywords

*Columba palumbus*; Azores; multiple regression models; abundance models; habitat selection

# Index

# List of tables

# List of figures

# 1. Introduction

## 1.1. Island biodiversity

Islands harbour nearly 17% (over 1750 species) of the extant world's bird species, of which about 23% are threatened, representing 39% of the threatened birds worldwide (Johnson and Stattersfield, 1990). The size, shape and degree of isolation make islands ecologically unique, but these characteristics are also responsible for the fragility and vulnerability of insular ecosystems (Gillespie and Clague, 2009). Islands have the highest proportion of recorded species extinction. It is estimated that, over the past 400 years, 90% of bird species extinctions occurred on islands (Johnson and Stattersfield, 1990). Islands are also known as centres of range restricted species and thus are areas with high levels of endemism (Kier et al., 2009).

Islands are frequently colonized by species from the closest mainland areas and due to their distinct environmental conditions those species can evolve under different selective pressures (Cox et al., 2005; Molles, 2013). This different evolutionary path will probably drive to disparate morphological characteristics of colonizers when compared with individuals of continental populations (Valen, 1973; Andrade et al., 2015). It has been suggested that patterns of morphological variation on islands can follow some predictable trends, known as "Island Rule" (Van Valen, 1973; Lomolino, 2005; Andrade et al., 2015). Briefly, smaller species have the tendency to become bigger and larger species tend to become smaller than their continental counterparts.

The Azores are an example of islands that exhibit endemic species and subspecies (Gillespie and Clague, 2009). It is the northernmost archipelago of Macaronesia (Figure 1.1), comprising nine islands of volcanic origin that emerged during the Tertiary and Quaternary periods (Whittaker and Fernández-Palacios, 2007). The animal and plant colonisations may have happened in the late Miocene and Pliocene and now, these islands are considered as displaying relatively high levels of genetic diversification, endemism and speciation compared to other oceanic islands (Adler et al., 1995; Francisco-Ortega et al., 2000; Juan et al., 2000; Dourado et al., 2013). The impact of direct habitat destruction of humans on islands is usually devastating. Many of these places were covered by native forests, but now those habitats face a partial or entire loss, many times by conversion to non-native forest, agriculture fields, pastures, and urban areas. Invasive species are a great threat to island biodiversity and they can have either indirect impacts on biodiversity through habitat alteration or direct impacts

through interaction with the native species (Paulay, 1994). The Azores were discovered in 1432, and over the last 500 years of human occupation, most of the original native forest, Laurissilva, was destroyed (Dias and Melo, 2010; Connor et al., 2012; Elias et al., 2016; Rull et al., 2017). In total, 420 endemic species (of a total of 4467) of fauna and flora were known to occur in the archipelago (Gillespie and Clague, 2009). Today, approximately 70% of the vascular plants and 58% of the arthropods found in Azores are exotic, and many of them are considered invasive (Gillespie and Clague, 2009). Several extinctions were already documented in Azores, for instance palaeoecological data indicate that at least two plant species on Pico (Connor et al., 2012) and the São Miguel Scops Owl (*Otus frutuosoi*) (Rando et al., 2013) went extinct probably due to human arrival and subsequent habitat alterations (Connor et al., 2012; Rando et al., 2013).


Figure 1.1 – The Azores archipelago location relatively to continental Europe and Africa.

## 1.2. The woodpigeon (*Columba palumbus*)

The woodpigeon (*Columba palumbus*) is a member of the family Columbidae (Cramp, 1985), that can be found in the western Palearctic, mainly in Europe and North Africa (Figure 1.2). Beyond the western Palearctic, its distribution extends slightly east of the Urals, and from the mountains of west-central Asia to southern Oman (Cramp, 1985). Essentially, it is a sedentary species, but the northern populations are migratory and winter in the southern part of the breeding range (Vaurie, 1961). The Iberian Peninsula is recognized as having an important role as wintering ground for Fennoscandian and

eastern European populations, which arrive in large numbers to the region from October onwards (Bea et al., 2003). In Europe, the subspecies *C. p. palumbus* breeds across the continent, from Russia and Scandinavia to the Atlantic coasts and to the Peninsulas across the northern Mediterranean Sea (Cramp, 1985). In mainland Portugal, the woodpigeon occurs throughout the country, being more common in northern areas, exhibiting a more continuous distribution in comparison to southern areas, where it is absent from some regions of Baixo Alentejo region (Eqipa Atlas, 2008; Catry et al., 2010). In the Azores islands occurs a distinct subspecies, the Azores woodpigeon (*C. p. azorica*), which is endemic to this archipelago (Dourado et al., 2013). Another endemic subspecies, the Madeiran woodpigeon (*C. p. madeirensis*), inhabited Madeira but it went extinct in the beginning of the 20th century (BirdLife International, 2016).

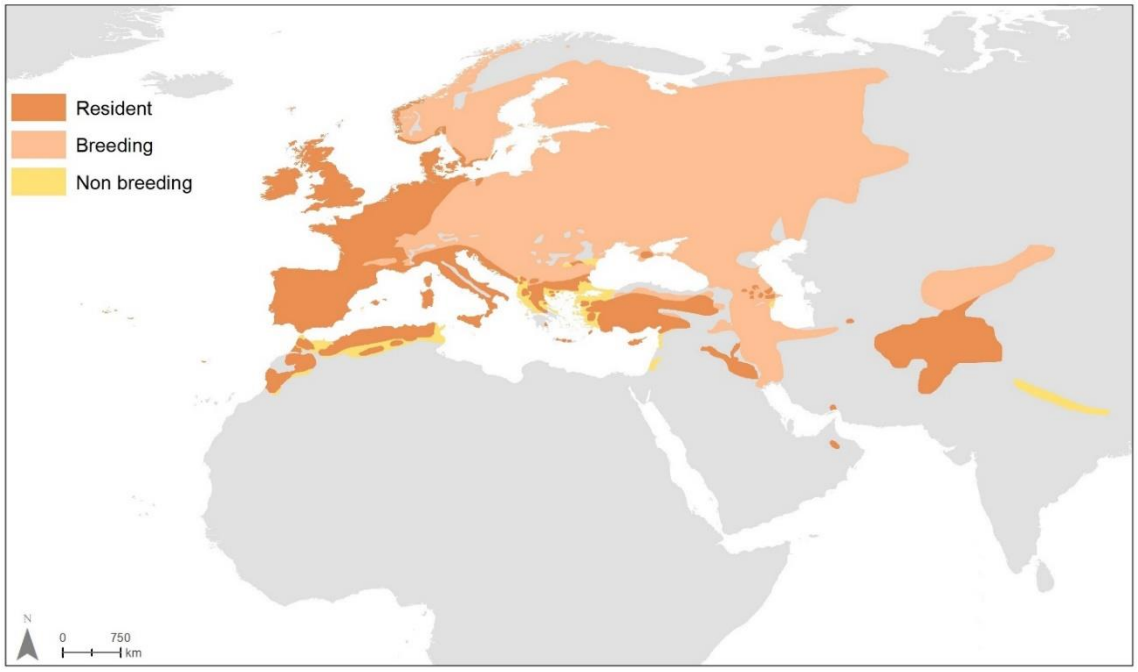



Figure 1.2 – Global distribution of the woodpigeon (*Columba palumbus*). (Adapted from BirdLife International and Handbook of the Birds of the World 2016).

According to Hartert and Ogilvie-Grant (1905), the Azores woodpigeon differs from the typical European woodpigeon in the deeper and more vinous colour of the chest and the darker slate-grey rump and head. Also, the longer upper tail feathers are more brownish and the under tail and wing coverts are generally darker and duller in comparison to continental individuals. The Azores woodpigeon has longer legs, shorter wings, narrower chest and thicker bill than continental specimens (Figure 1.3) (Andrade et al., 2016, Cataldo, 2017). The Azorean subspecies has only been recognized as separate subspecies based only on external characters (Andrade et al., 2016).

Dourado et al. (2013) found no genetic differentiation between the continental *C. p. palumbus* and the endemic *C. p. azorica*, in two mitochondrial and one nuclear gene, suggesting a relatively recent colonisation of the Azores islands. However, other study, using genome-wide single nucleotide polymorphism concluded that the Azores woodpigeon has already undergone through slight differentiation from the continental subspecies (Andrade et al., 2016).

Figure 1.3 – Left: Azores woodpigeon (*Columba palumbus azorica*); Madalena, Pico, Azores; May 9, 2017. Right: Woodpigeon (*C. p. palumbus*); Copenhagen, Denmark; January 11, 2011. (Photos by Tiago Rodrigues).

Currently, the European population of woodpigeon (41 to 58 million mature individuals) represents 80% of the global population (51 to 73 million mature individuals) (BirdLife International, 2016). The species is considered as “Least Concern” due to the extremely wide distribution area, the high population numbers and apparent increasing population trends (BirdLife International, 2016). In continental Portugal, the species is categorised with the same conservation status, whereas in the Azores is considered as “Data Deficient” (Cabral et al., 2005).

### 1.2.1. Ecology and behaviour of continental populations

The woodpigeon is an ecologically plastic species that occurs in a large range of climatic conditions, from the low-arctic, to Mediterranean regions (Cramp, 1985). It avoids wide open areas, such as bare rocky mountain regions, plains or uplands and exposed seacoasts, as well as densely vegetated wetlands. Apparently, the species thrives in habitats composed by patches of woodlands (deciduous or coniferous) alternated with open spaces (Cramp, 1985). In northern Portugal, the woodpigeon inhabits coniferous forests, eucalyptus forests with understory, oak forests, and mixed

shrubland near to agriculture fields, where southern population are associated to cork and holm oak forests with shrubland (Catry et al., 2010).

The duration of the breeding season across Europe varies considerably, even within countries, with urban birds nesting earlier than rural (Cramp, 1985). Overall it can start in February and end in November (Murton and Isaacson, 1962). In the beginning of the breeding season, males start marking territories by advertising-call and display-flight (Murton and Isaacson, 1962; Cramp, 1985). There are few studies quantifying home-range dynamics of the woodpigeon, but in Cambridgeshire, England, it remains close to the nests during the breeding season utilizing about 275 ha. In autumn, birds extend their range and the juveniles use larger areas (1283 ha) than the adults (628 ha) (Haynes et al., 2003).

In continental Europe, woodpigeon feeds mainly on plant material, as greens leaves, seeds, fruits, buds, flowers, root crops and occasionally invertebrates especially during the breeding season (Murton et al., 1964; Cramp, 1985). In the Iberian Peninsula, during winter, most of its diet is composed by holly oak and cork oak acorns and seeds of various cereals (Purroy and Purroy, 2016).

The woodpigeon is generally a gregarious species, especially when feeding and roosting outside the breeding season, forming at times flocks with thousands of birds (Cramp, 1985). During that time of the year, woodpigeons appear to form larger flocks during the morning when feeding, but show a tendency to break up in smaller groups in the afternoon, when returning to the roost (usually in tall trees) (Murton, 1965). In winter woodpigeons are quite mobile and may fly up to 65 km to feeding grounds (Cramp, 1985). A feeding flock does not occupy a whole field. Instead, the birds move all together as a well-defined group. During the breeding season, it shows a solitary and territorial behaviour, though several pairs sometimes nest close together, particularly in urban areas (Cramp, 1985).

### 1.2.2. The case of the Azorean population

In the beginning of the 20$^{th}$ century, the size of the Azores woodpigeon population declined supposedly as consequence of the natural woodlands conversion into pastures for dairy farm (Dickens and Neves, 2005). At that time, it was considered as little numerous, especially on Santa Maria, São Miguel, Graciosa, Terceira and Faial, and absent from Flores and Corvo. São Jorge and Pico were the islands with greater population numbers (Hartert and Ogilvie-Grant, 1905; Bannerman and Bannerman,

1966). It is currently considered common, especially on Pico, Graciosa and Terceira islands, and less common on Flores and Corvo (Lucas, 2004; Dickens and Neves, 2005; Fontaine et al., 2014). The information about its distribution in Azores is available in the Portuguese atlas of breeding birds (Figure 1.4) (Equipa Atlas, 2008). However, the mapping scale of the distribution (10×10 km) is too coarse to have a precise idea where the woodpigeon occurs in the archipelago and few studies have been done on the environmental and ecological factors that may be related with the observed distribution of Azores woodpigeon abundance (Lucas, 2004). The Azores woodpigeon abundance seems to be higher in agriculture fields, at least during winter (Lucas, 2004), but the species can also be found in exotic plant formations, as *Cryptomeria japonica* forests, and in natural vegetation (e.g. *Morella faya*, heather, *Juniperus brevifolia, Picconia azorica*) (Lucas, 2004; Equipa Atlas, 2008). During winter, the distribution of the abundance throughout the habitats does not seem to change along the day (Lucas, 2004), but there is no information for other seasons.

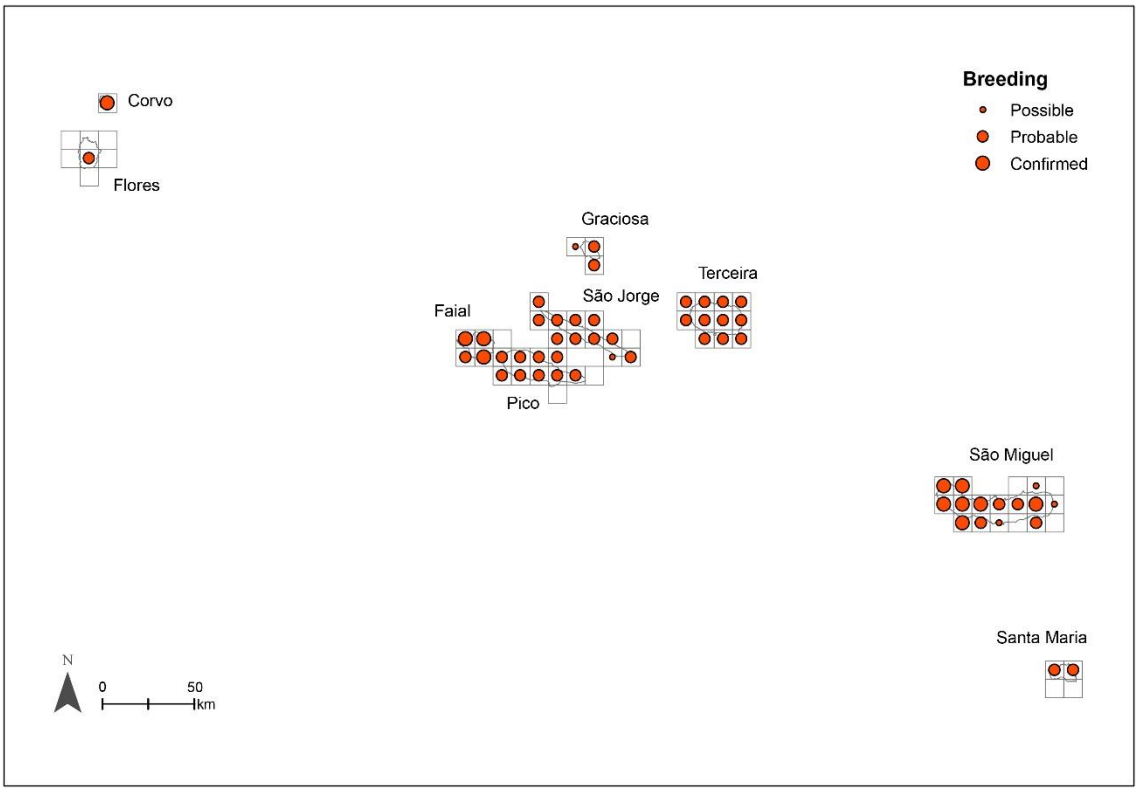



Figure 1.4 – Distribution of the Azores woodpigeon (*Columba palumbus azorica*). Adapted from Equipa Atlas (2008).

It breeds in all islands of the archipelago (Equipa Atlas, 2008), however, little is known about its reproductive biology (Fontaine et al., unpublished). According to Fontaine et al. (unpublished), 56% of the nests found on Terceira were built in three exotic species of trees. Similar observations were made in other islands, suggesting that the Azores

woodpigeon is not dependent on natural vegetation for breeding and that it exhibits high adaptation levels, including to urban areas (Fontaine et al., unpublished). The breeding season ranges from late February to early October, with similar duration as observed in continental populations (Fontaine et al., unpublished). Usually one or two eggs are laid (Fontaine et al., unpublished). There is little knowledge about the diet, but it is likely that the Azores woodpigeon feeds mainly on plant material, such as crop seeds and fruits from natural vegetation (Hartert and Ogilvie-Grant, 1905; Lucas, 2004).

### 1.2.3. Interactions between pigeons and humans

The woodpigeon is frequently blamed for damage on crops in Europe (Murton, 1965). In the United Kingdom was considered the major bird pest feeding on a wide range of arable crops (Inglis et al., 1994). However, these cases seem to be rare and the effects of a bird attack may be difficult to measure because many factors can be operating at the attack time or latter (Murton, 1965). The Azorean populations apparently behave similarly to continental populations, as local farmers report that woodpigeons often frequent agricultural fields, being pointed for damages to crops and vineyards on Pico and Terceira islands (Fontaine et al., unpublished; Fontaine et al., 2014). The negative impact seem to be stronger on Terceira, which might be related to the relatively high abundance of woodpigeons on the island (Fontaine et al., unpublished). There are many examples of wild animals responsible for damaging crops on islands. In some occasions, the extension of farming in tropical and subtropical regions is so wide that wild animals, including endemics, become farm pests (Marrero et al., 2004). In Madeira, the Madeira laurel pigeon (*Columba trocaz*) is considered the major bird pest of the archipelago (Oliveira et al., 1999, 2006), visiting regularly agriculture lands close to forest patches, feeding on a great variety of crops and fruits (Marrero et al., 2004). In the Mentawai islands, the Pagai Island macaque (*Macaca pagensis*) is blamed for damage gardens and sweet potatoes culture (Lee and Priston, 2005). To minimise conflicts between humans and the Azores woodpigeon it is crucial to have knowledge about local scale distribution of the species and habitat selection patterns, and how these change throughout the year and throughout the day. Studying a relevant ecological state variable, as the spatial distribution pattern of abundance (Kéry et al., 2005), is an informative way to assess the preferences of a species across the environment (Rovero et al., 2014).

## 1.3. Regression models to estimate population abundance

Regression models rely on regression analysis, which is a statistical process for estimating the relationships among variables (Nally, 2000). They are often used in ecology to determine patterns affecting species and to estimate population abundance, population size and species richness (Nally, 2000; Lobo et al., 2002; Tognelli and Kelt, 2004; Lyashevska et al., 2016).

Abundance, defined as the number of individuals registered per sample unit, is a variable of a fundamental importance in ecology studies, especially in monitoring programs (Kéry et al., 2005). In studies of this nature, abundance data can provide advantages over presence-absence and presence-only data, because it is known to be a more detailed form of data (Pearce and Ferrier, 2001; Howard et al., 2014; Lyashevska et al., 2016). Indexes of relative abundance and density are likely to be a good indicator of habitat quality, reflecting key demographic and environmental factors, such as carrying capacity, reproductive success, longevity and susceptibility of populations to extinction (Pearce and Ferrier, 2001; Howard et al., 2014). Furthermore, abundance seems to enable a greater ability to discriminate species' range boundaries, because it is better related to conservation status, extinction risk, community structure and function and suitability of the habitats than presence-absence data (Howard et al., 2014). It also improves the predictive mapping performance of habitat quality and consequently of the species under study (Pearce and Ferrier, 2001; Howard et al., 2014). According to Howard (2014), models built using abundance data generally outperform the ones built with presence-absence data, being more accurate than those. Although a positive correlation between habitat quality and species abundance is assumed, interactions between species and habitat characteristics and effectiveness of sampling techniques adds some uncertainty to this assumption. Even with an effective sampling method, the detectability of a species will also play a role here and will influence recorded abundance or density (Pearce and Ferrier, 2001). For some taxa, the uncertainty of abundance data might be much higher than those of incidence data, which makes abundance maps more uncertain (Lyashevska et al., 2016).

Many factors influence species' distributions and ideally, all of them should be included in the model. However it is difficult to do so, because these are complex interactions and, in most cases, there is no sufficient information to include in models (Anderson et al., 2002). For example, competition and other biotic interactions are many times neglected resulting in an underestimation of the realized niche (Guisan and Thuiller, 2005). It was already shown that the inclusion of additional predictor variables of

competitive species can significantly increase the predicting power of models (Leathwick and Austin, 2001). Both species and environmental data are usually sampled in a limited time and space, thus models built using these predictors can only reflect a snapshot view of the expected relationship. Knowing this, one must assume that the modelled species is in pseudo-equilibrium with its environment and time (Guisan and Thuiller, 2005). Another common problem when building models is identifying the appropriate scale for modelling (Guisan and Thuiller, 2005). That is, one may address independently the ideal resolution (grain size) and extent of the study area (Wiens, 2002; Guisan and Thuiller, 2005), because these components, at which data is measured or analysed, are crucial to achieve the best results possible (Elith and Leathwick, 2009). Mismatches can occur between the resolution at which the species data were sampled and the one at which environmental predictors are available. Understanding the relationship between the species and the processes responsible for the observed distribution patterns is essential to avoid a mismatch between the scale used for modelling and the one at key processes occur, because patterns observed on one scale may not be apparent on another scale (Guisan and Thuiller, 2005). Contradictory information to the same ecological question can be provided by the use of different geographical extents. Similar reasoning holds for resolution, once the same environmental parameter sampled in distinct resolutions may have very different meanings for a given species (Guisan and Thuiller, 2005). Finally, the right scale will also depend on the biological traits of the species under study. On one hand, highly mobile organisms might require the use of large-sized modelling cells, in a way that all types of habitats can be included and fulfil the different requirements of the species. On the other hand, for less mobile organisms, small-sized cells should be used to provide better predictions (Guisan and Thuiller, 2005). In another words, the appropriate scale is dictated by the study goals, the available environmental data, and the species under study (Elith and Leathwick, 2009).

When only environmental predictors are used to build a model and the geographic predictors (two-dimensional map coordinates or three-dimensional digital elevation models) are left apart, what is being modeled in reality is the variation in occurrence or abundance of a species in environmental space (Elith and Leathwick, 2009). Such model does not reflect geographic proximity even when the predictors are mapped into geographic space. Ultimately species distribution models with solely environmental predictors simply reflect the spatial autocorrelation of the environmental predictors (Elith and Leathwick, 2009).

The quality of regression-based models is highly reliant on the number of observations of the dependent variable (presence-absence or abundance) used in the model building process (Wisz et al., 2008). Several studies have shown that model performance increases while variability in predictive accuracy decreases with increasing sample size (Carroll and Pearson, 1998; Hernandez et al., 2006; Wisz et al., 2008). There are three main reasons that explain why model performance usually increases with sample size: i) bigger sample sizes usually translates in lower levels of uncertainty associated with parameter estimates; ii) outliers tend to have a higher weight in analysis when the sample sizes are smaller; iii) ecological niches of species are very complex systems and larger sample sizes may be important to achieve an accurate description of the range of conditions over which a species occurs (Wisz et al., 2008). It is also important to refer that the sample size should be proportional to the study area size, rather than chosen indefinitely independently of the dimensions of the study area (Hernandez et al., 2006).

Various families of models can be used in species distribution models, although generalized linear models and generalized additive models are the more common and have been widely used since the late 1980s (Franklin, 2010). More recently, multivariate adaptive regression splines started to be used (Franklin, 2010).

## 1.4 Conceptual framework of the thesis

Knowledge about the distribution of the Azores woodpigeon abundance at local-scale and its habitat selection patterns are still limited. It is important to understand and predict its movements according to daily and seasonal variations for the management and conservation of its populations. The proposal is to apply models of relative abundance as they are known to provide informative indicators about habitat quality that reflect the reproductive success, longevity, carrying capacity, and susceptibility of population to extinction (Pearce and Ferrier, 2001). Beyond the topics regarding the woodpigeon ecology addressed in this work, it is also important to investigate whether abundance is a meaningful variable for ecological modelling when studying species that are very abundant and widespread.

## 1.5 Objectives

The main goal of this study was to investigate patterns in the distribution of abundance and habitat use of the Azores woodpigeon. More specifically, we aimed to: i) investigate seasonal and daily variations in habitat use of the Azores woodpigeon on Terceira island, to understand the movement dynamics of the species; ii) map the distribution of abundance and identify environmental factors that may be related with observed distribution of woodpigeon abundance; iii) model the potential distribution of woodpigeon abundance and compare the results of three different measures of abundance.

We expected the woodpigeon to be more abundant in regions with potential greater food availability, such as in the agriculture fields and natural vegetation areas and less abundant in regions with exotic plant species. However, as it relies on trees for nesting, during the breeding season the abundance should increase in forested areas. We also predicted that the abundance should be a good state variable to build the ecological models, providing a good insight of the distribution and habitat use of the Azores woodpigeon.

# 2. Methods

## 2.1. Study area

The study area was the Terceira island (between latitude 38°38'21.1"N to 38°48'12.3"N and longitude 27°02'28.5"W to 27°22'45.9"W), one of the largest islands of the Azores archipelago, located in the central group, with 402.2 km$^2$ and with its highest point at 1023m (Figure 2.1) (Borges et al., 1999). The climate of the island is of subtropical nature (Gillespie and Clague, 2009). Laurel forests dominated in the past, but have declined with the increasing human activities that replaced it by agriculture fields (Gillespie and Clague, 2009). Compared to other islands, Terceira present relatively high values of woodpigeons abundance and many farmers blame them for damaging crops (Dickens and Neves, 2005).

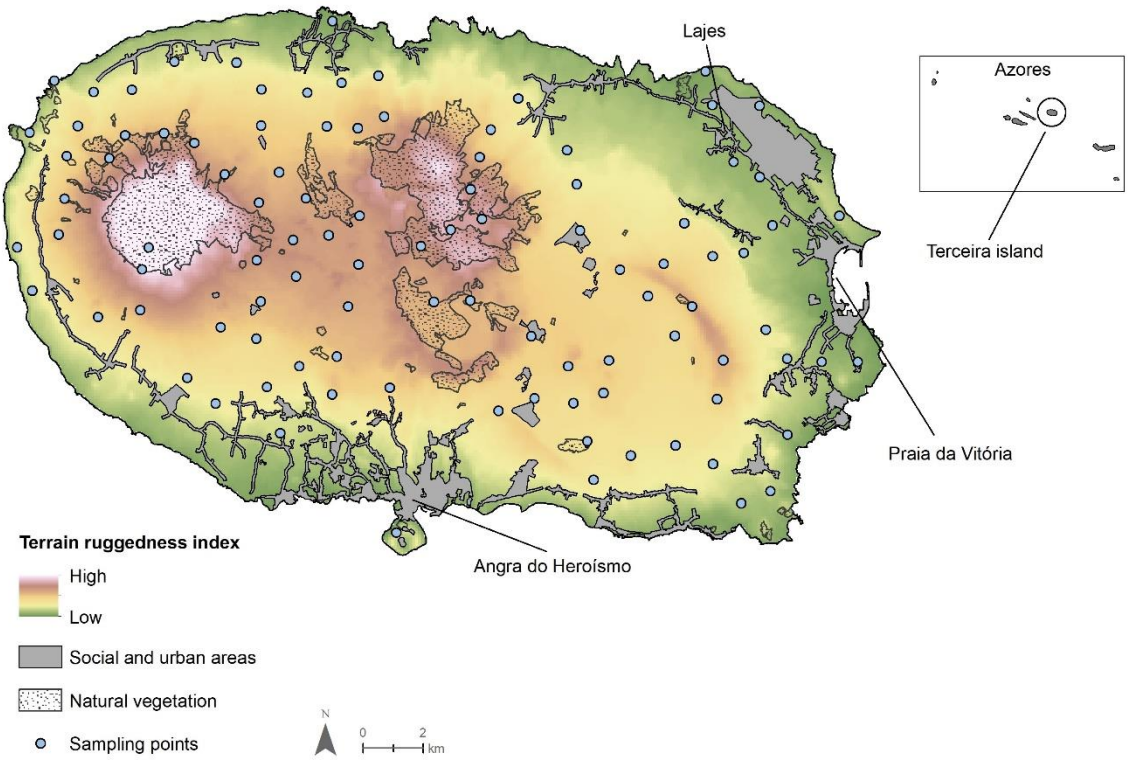



Figure 2.1 –Terceira island map, displaying Terrain ruggedness index, social and urban areas, natural vegetation, and the distribution of sampling points used in this study. Location of Terceira in the Azores archipelago.

## 2.2. Data availability

For this study we used a dataset on Azores woodpigeon abundance collected on Terceira, from October 2012 to August 2014, during the field work for the project AZORPI (M2.1.2/I/025/2011), funded by Regional Government of Azores (namely the *Direção Regional da Ciência e Tecnologia*). Data were collected on 100 sampling points (Figure 2.1). These points were defined randomly in ArcGIS 10.1 (ESRI, 2011) with spatial coordinates associated to each one of them. They were located at least two km apart from each other, which reduced the probability of double counts (e.g. counting the same individual in different points the same day). The minimum distance used was chosen according to the territorial behaviour and home range size of the species (Murton, 1965). To facilitate the logistics (displacement, time, and costs) and to further decrease the probability of double counts, "groups" of points were visited in sequence, the same day, and the observer travelled by car between points. The Nearest Neighbour index, measuring the clustering level of the sampling points, informed that points were spatially dispersed (nearest neighbour ratio = 1.188; Z-score = 3.604; $p < 0.001$; ANNEX 1).

The sampling points were monitored fortnightly from October 2012 to August 2014, during mornings (from sunrise until three hours after) and afternoons (from three hours before sunset) (Table 1). At each sampling point, the observer registered each Azores woodpigeon observation (individuals seen or listen) during 10 minutes, namely the number of individuals, age, behaviour (whether the animal was flying, perched or landing), plus the time of each observation, weather conditions (percentage of cloudiness, wind according to Beaufort's scale and whether it was raining or not). The sampling period (10 minutes) in each point started immediately after the arrival to the point (i.e. without waiting period, Lee and Marsden, 2008), to count birds that could be scared away by the observer. All the birds that took flight at the arrival to the point, within a radius of 250 m, were considered. Surveys were only conducted in good visibility conditions.

October 2012 was considered a training month and thus was not considered in the analyses. In the remaining sampling period, a total of 2690 observations of pigeons perched or landing were registered, of which 1395 during the morning period and 1295 during the afternoon. The dataset was divided in two years: Year 1, including data collected from November 2012 to September 2013; and Year 2, including data collected from October 2013 to August 2014. A total of 1386 and 1304 observations were collected in the first and second years, respectively. The dataset was also divided

into Breeding season, from March to September, and a Non-breeding season, from October to February, according to the breeding ecology of the Azores woodpigeon (Fontaine et al., unpublished). In the Breeding season 1859 observations were accounted, while in the Non-breeding season 831 observation were accounted.

Abundance (number of pigeons / 10 minutes of observation) was quantified for each point. For this, only landing and perched individuals were included; birds flying were excluded because they were not associated to any specific habitat. For analyses, three datasets were generated, for each point: the average abundance including all data (hereafter *All data*), the average abundance excluding zero counts (*Truncated data*), and the maximum abundance ever recorded (*Maximum data*); *All data* abundance estimates followed a zero-inflated Poisson distribution in both seasons, whereas the Truncated and *Maximum data* had a Poisson distribution (Figure 2.2).

Table 1 – Sampling periods starting in November of 2012 and ending in August of 2014. October of the Year 1 was considered a training month. NB = Non-breeding season; B = Breeding season. Start and end dates (day/month) for each visit.

| Year | Season | Month | Visit | Start | End | Visit | Start | End |
|---|---|---|---|---|---|---|---|---|
| Year 1 | NB | Oct | 1 | 11/10 | 19/10 | | | |
| | | Nov | 2 | 28/10 | 08/11 | 3 | 18/11 | 26/11 |
| | | Dec | 4 | 02/12 | 17/12 | 5 | 25/12 | 05/01 |
| | | Jan | 6 | 06/01 | 18/01 | 7 | 20/01 | 29/01 |
| | | Feb | 8 | 03/02 | 13/02 | 9 | 17/02 | 28/02 |
| | B | Mar | 10 | 04/03 | 20/03 | 11 | 24/03 | 03/04 |
| | | Apr | 12 | 06/04 | 17/04 | 13 | 19/04 | 03/05 |
| | | May | 14 | 05/05 | 13/05 | 15 | 18/05 | 28/05 |
| | | Jun | 16 | 03/06 | 12/06 | 17 | 19/06 | 25/06 |
| | | Jul | 18 | 01/07 | 08/07 | 19 | 28/07 | 08/08 |
| | | Aug | 20 | 11/08 | 16/08 | 21 | 25/08 | 31/08 |
| | | Sept | 22 | 06/09 | 12/09 | 23 | 23/09 | 04/10 |
| Year 2 | NB | Oct | 24 | 09/10 | 17/10 | 25 | 21/10 | 04/11 |
| | | Nov | 26 | 08/11 | 13/11 | 27 | 23/11 | 28/11 |
| | | Dec | 28 | 07/12 | 14/12 | 29 | 20/12 | 30/12 |
| | | Jan | 30 | 08/01 | 14/01 | 31 | 23/01 | 28/01 |
| | | Feb | 32 | 08/02 | 08/02 | 33 | 16/02 | 24/02 |
| | B | Mar | 34 | 08/03 | 13/03 | 35 | 20/03 | 27/03 |
| | | Apr | 36 | 06/04 | 11/04 | 37 | 17/04 | 25/04 |
| | | May | 38 | 02/05 | 06/05 | | | |
| | | Jun | 39 | 07/06 | 16/06 | 40 | 22/06 | 26/06 |
| | | Jul | 41 | 02/07 | 11/07 | 42 | 21/07 | 30/07 |
| | | Aug | 43 | 09/08 | 14/08 | 44 | 22/08 | 30/08 |

## 2.3. Daily, seasonal, and yearly variations in abundance

We investigated differences in the distribution of the average abundance estimates according to putative daily variations (Morning and Afternoon), seasonal variations (Breeding season and Non-breeding season) and yearly variations (Year 1 and Year 2). Five scenarios were tested: i) daily variations (Morning vs. Afternoon); ii) seasonal variations (Breeding vs. Non-breeding), iii) daily variations during the Breeding season; and iv) daily variations during the Non-breeding season; v) yearly variations (Year 1 vs. Year 2). Analyses were performed using dispersion graphs, Spearman's rank correlation coefficient and ANOVAs for each scenario.

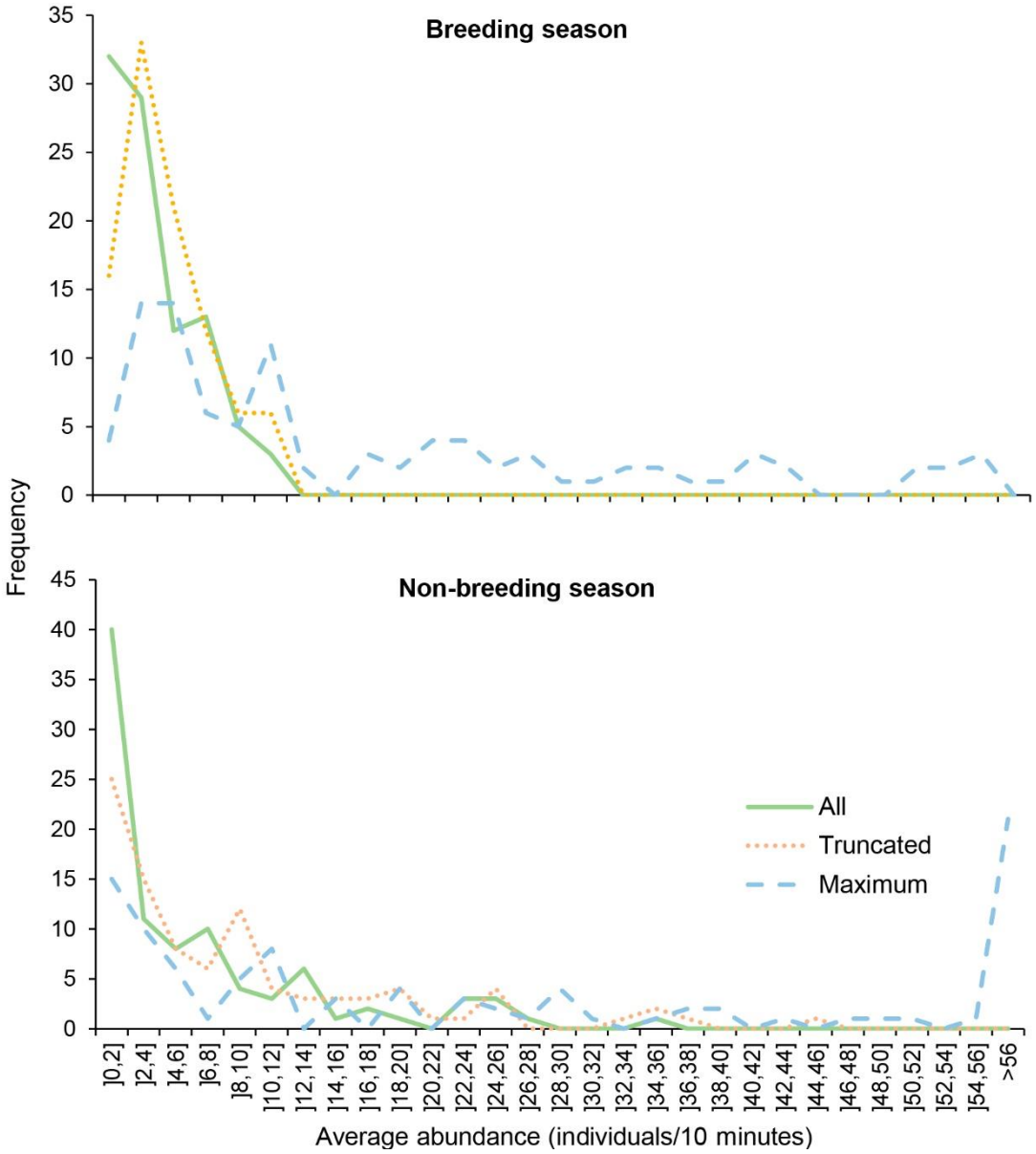



Figure 2.2 – Frequencies of abundance values (number of pigeons / 10 minutes of observation) including *All data* (solid green line), *Truncated data* (dotted orange line) and *Maximum data* (dashed blue line) for the Breeding season (top) and the Non-breeding season (bottom).

## 2.4. Mapping the distribution and abundance

We mapped the distribution of the Azores woodpigeon on Terceira using the three types of abundance estimates. In addition to the distribution of abundance, we also included the respective standard variation. Maps of abundance were only produced for distinct periods (day, season, or year) when significant abundance differences were recorded. The maps were built in ArcGIS 10.1 (ESRI, 2011).

## 2.5. Identification of environmental factors related with abundance

Three types of environmental factors were considered to relate with the estimated abundances: habitat, topography and climate. Habitat variables included nine habitat categories extracted from the "Forest Inventory" (IFRAA, 2007) and the hydrographic network extracted from military maps (Table 2, ANNEX 2). The nine habitat categories were chosen according to the total area occupied by them and according to the ecological relevance they could have. Habitat variables (Table 2) were rasterized to a pixel size of 2×2 m and then upscaled to 100×100 m. The upscaled pixel quantified the percentage of coverage of each habitat category and hydrographic network. This step allowed the maintenance of the spatial detail of variables with low coverage, as the case of watercourses.

Table 2 – Habitat categories (HC) and individual habitat levels (HL) and their correspondent occupied areas and percentages (IFRAA, 2007).

| HC Code | HC description | HL description | Area HC ($m^2$) | % HC | Area HL ($m^2$) | % HL |
|---|---|---|---|---|---|---|
| AGR | Agriculture fields | Agriculture fields or pastures | 26479.7 | 66.1 | 26479.7 | 66.1 |
| CON | Coniferous forests | *Pinus pinaster* | 104.3 | 0.3 | 129.2 | 0.3 |
| | | Other resinous trees | 19.3 | 0.0 | | |
| | | *Chamaecyparis lawsoniana* | 5.6 | 0.0 | | |
| CRYP | *Cryptomeria* plantations | *Cryptomeria japonica* | 1609.9 | 4.0 | 1609.9 | 4.0 |
| EUC | Eucalyptus plantations | *Eucalyptus globulus* | 2575.0 | 6.4 | 2575.0 | 6.4 |
| EXO | Other exotic forests | *Pittosporum undulatum* | 1347.9 | 3.4 | 1570.1 | 3.9 |
| | | *Acacia melanoxylon* | 174.2 | 0.4 | | |
| | | Other broadleaves | 48.0 | 0.1 | | |
| NAT | Natural vegetation | Natural Vegetation | 3769.1 | 9.4 | 3798.5 | 9.5 |
| | | *Myrica faya* | 29.4 | 0.1 | | |
| SU | Social and urban areas | Urban aggregations | 2066.4 | 5.2 | 2987.2 | 7.5 |
| | | Social areas | 920.8 | 2.3 | | |
| UNCUL | Uncultivated areas | Uncultivated land | 469.0 | 1.2 | 469.0 | 1.2 |
| UNPRO | Unproductive areas | Unproductive land | 422.5 | 1.1 | 422.5 | 1.1 |

The altimetry was available in contour and was converted in Terrain ruggedness index (TRI) at a pixel size of 100×100 m in ArcGIS.

Climatic variables, available at a 100×100 m pixel size, included estimates of monthly average temperature, accumulated precipitation, and relative humidity from 30 years (Azevedo, 1996). A total of 12 variables were constructed for each of the three climatic traits, each one corresponding to a month of the year.

The final dataset included 96 normalized variables (Table 3). Correlations between variables were tested using Pearson correlation test in R 3.2.4 (R Development core Team, 2016) using the Hmisc R package (Harrell, 2017) and considered significant at $p < 0.05$.

Table 3 – Variables considered for models, their descriptions and units. Variables marked * include 12 variables, one corresponding to each month of the year.

| Code | Description | Units |
|---|---|---|
| AGR | Land cover of agriculture fields | % |
| CON | Land cover of coniferous forests | % |
| CRYP | Land cover of *Cryptomeria japonica* plantations | % |
| EUC | Land cover of eucalyptus plantations | % |
| EXO | Land cover of exotic forests | % |
| NAT | Land cover of natural vegetation | % |
| SU | Land cover of social and urban areas | % |
| UNCUL | Land cover of uncultivated areas | % |
| UNPRO | Land cover of unproductive areas | % |
| HABHET | Habitat heterogeneity | Adimensional |
| HYDRO | Land cover of water courses (hydrography) | m |
| TRI | Terrain ruggedness index | Adimensional |
| TMAX* | Maximum average temperature | °C |
| TMED* | Average temperature | °C |
| TMIN* | Minimum average temperature | °C |
| PREC* | Accumulated rainfall | mm |
| HMAX* | Maximum average relative humidity | % |
| HMED* | Average relative humidity | % |
| HMIN* | Minimum average relative humidity | % |

A Principal Components Analysis (PCA) was done using ArcGIS to overcome the high correlations found between variables. The first ten Principal Components (PCs), accumulating 98% of the variance among the original variables (Table 4), were retained for further analyses. The spatial representation of the first ten PCs is characterised in ANNEX 3.

Finally, a buffer of 500 m of radius around each sampling point was created, and the mean of the PCs within the buffer was calculated and extracted to perform the

modelling. This distance corresponds to the maximum distance at which pigeons were registered from the sampling points.

Table 4 – Environmental factors with highest contributions to the first 10 Principal Components (PC) estimated by Principal Component Analysis. The spatial representation of PCs is depicted in ANNEX 3. The variables that most contribute on each of the three components retained by the models (PC01, PC02 and PC03) are highlighted in grey.

| Variables | PC01 | PC02 | PC03 | PC04 | PC05 | PC06 | PC07 | PC08 | PC09 | PC10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Eigenvalues (EV) | 5155.9 | 1571.6 | 449.7 | 369.2 | 272.5 | 206.1 | 188.8 | 161.4 | 88.1 | 70.7 |
| % EV | 59.2 | 18.0 | 5.1 | 4.2 | 3.1 | 2.3 | 2.1 | 1.8 | 1.0 | 0.8 |
| Cumulative EV | 59.2 | 77.2 | 82.4 | 86.6 | 89.7 | 92.1 | 94.3 | 96.1 | 97.1 | 98.0 |
| AGR | 0.000 | -0.009 | -0.006 | -0.003 | -0.015 | 0.000 | -0.004 | -0.005 | 0.037 | -0.016 |
| CON | 0.171 | 0.888 | 0.070 | -0.017 | 0.112 | -0.020 | 0.021 | -0.041 | 0.000 | 0.188 |
| CRYP | -0.056 | -0.064 | -0.173 | 0.299 | -0.402 | 0.585 | -0.406 | 0.083 | -0.204 | 0.127 |
| EUC | -0.005 | -0.187 | -0.662 | -0.319 | 0.496 | 0.005 | 0.045 | 0.110 | -0.117 | 0.147 |
| EXO | 0.013 | -0.112 | -0.097 | -0.049 | -0.592 | -0.415 | 0.494 | -0.032 | -0.252 | 0.107 |
| NAT | -0.203 | -0.204 | 0.618 | -0.567 | 0.077 | 0.138 | -0.087 | -0.035 | -0.028 | 0.170 |
| SU | 0.077 | -0.255 | 0.273 | 0.632 | 0.402 | -0.314 | -0.085 | -0.170 | -0.128 | 0.129 |
| UNCUL | -0.005 | -0.024 | -0.025 | 0.020 | -0.040 | -0.021 | -0.007 | 0.024 | 0.488 | -0.702 |
| HABHET | 0.006 | -0.122 | -0.083 | 0.094 | -0.105 | 0.061 | 0.123 | -0.092 | 0.763 | 0.584 |
| HYDRO | -0.010 | -0.009 | -0.068 | 0.005 | 0.081 | 0.390 | 0.349 | -0.83 | -0.087 | -0.124 |
| TRI | -0.237 | 0.037 | -0.019 | 0.004 | 0.007 | -0.035 | -0.034 | -0.044 | -0.017 | 0.044 |
| RAIN_JAN | -0.197 | 0.049 | 0.015 | 0.110 | 0.074 | 0.080 | 0.214 | 0.156 | 0.009 | -0.022 |
| RAIN_FEB | -0.192 | 0.044 | 0.020 | 0.104 | 0.064 | 0.083 | 0.195 | 0.142 | 0.003 | -0.008 |
| RAIN_MAR | -0.201 | 0.041 | 0.016 | 0.094 | 0.059 | 0.078 | 0.184 | 0.133 | 0.002 | -0.013 |
| RAIN_APR | -0.191 | 0.034 | 0.020 | 0.084 | 0.051 | 0.080 | 0.167 | 0.117 | -0.006 | -0.003 |
| RAIN_SEP | -0.190 | 0.023 | 0.025 | 0.047 | 0.023 | 0.050 | 0.069 | 0.041 | -0.015 | 0.002 |
| RAIN_OCT | -0.202 | 0.033 | 0.024 | 0.075 | 0.040 | 0.070 | 0.132 | 0.092 | -0.006 | -0.004 |
| RAIN_NOV | -0.196 | 0.036 | 0.020 | 0.082 | 0.047 | 0.074 | 0.152 | 0.105 | -0.003 | -0.003 |
| RAIN_DEC | -0.198 | 0.045 | 0.017 | 0.107 | 0.070 | 0.086 | 0.213 | 0.157 | 0.004 | -0.017 |

## 2.6. Modelling the distribution of abundance

Relationships between the estimated abundances (*All data*, *Truncated data*, and *Maximum data*) and environmental factors (the 10 first PCs) were tested with multiple regression models for two distinct periods, the Breeding season and the Non-breeding season. Models were selected according to their Akaike's information criterion value. We used the best model to determine the importance of predictors and their significance for each model. During preliminary data inspections, six sampling points were considered as outliers and were removed from subsequent analyses. As previously referred the distributions of abundance are non-normal. Transformations are often used to transform a non-normal dependent variable into a normal one since "normality" is an important assumption for many statistical techniques (Box and Cox, 1964; Fraser, 1967; Andrews, 1971; Atkinson, 1973; Carroll and Ruppert, 1981).

Acknowledging this we tried five different transformations: Box-Cox $x(\lambda) = \frac{x^{\lambda}-1}{\lambda}$, reciprocal transformation ($1/x$), cube root ($\sqrt[3]{x}$), square root ($\sqrt{x}$) and square root of the square root ($\sqrt{\sqrt{x}}$).

A total of 10 model replicates were built for each of the three types of abundance estimations to account for variance within the dataset. In each model replicate, abundance data were randomly partitioned into training (79 sampling points) and test (15 points) datasets. Models were produced using 3.2.4 (R Development core Team, 2016) MuMIn R package (Barton, 2015).

For an easier interpretation of the results, after modelling all the transformations were reversed.

The signals of the coefficients of the PCs were analysed to understand if relationships between observed abundances and important variables were positive or negative. Plots depicting the variation of abundance in relation to the environmental variation of the most important variables were produced to visualise those relationships. These analyses allowed understanding how environmental variation can be associated with the spatial variation of abundance in distinct seasons and how abundance varies in relation to the variation along the environmental gradients.

# 3. Results

## 3.1. Daily, seasonal, and yearly variations in abundance

In all scenarios examined there was some correlation between temporal variations (minimum $R^2$ was 0.312; Table 5, Figure 3.1). The *Maximum data* abundance showed the lowest $R^2$ values, whereas the *All data* and *Truncated data* abundances exhibited the highest $R^2$. Regarding the Spearman rank correlation tests (Table 5), *All data* abundance once again had the highest values and the *Maximum data* the lowest ones, except for the Morning against Afternoon and for the Mornings compared with Afternoons during the Non-breeding season, being the *Maximum data* values similar or slightly higher than the *Truncated data*, respectively. Analyses of variance suggest that the distribution of abundances in the Breeding season and Non-breeding season are distinct (Table 5). The differences between the Breeding and the Non-breeding season were most noticeable in the *Truncated data* (p = 0.004) and *Maximum data* (p = 0.022) than in *All data* (p = 0.112). Besides this scenario, no other exhibited statistically significant differences.

Table 5 – Statistical tests (R2, Spearman rank correlation test and ANOVA) concerning variation in abundance among observation points for the tested scenarios, considering each dataset of abundance (All – all data, Truncated – excluding zero counts; Maximum – maximum abundance ever recorded): i) daily variation; ii) seasonal variations, iii) daily variation during the breeding season; and iv) daily variation during the non-breeding season; v) yearly variation. The bold font represents the statistical significant p-values (p < 0.05).

| Scenarios | Variable | $R^2$ | Spearman rank correlation | | ANOVA | | | |
|---|---|---|---|---|---|---|---|---|
| | | | rho | p-value | df | Error | F | p-value |
| i | All | 0.486 | 0.800 | < 0.001 | 197 | 7565.3 | 0.000 | 0.985 |
| | Truncated | 0.493 | 0.740 | < 0.001 | 195 | 8600.2 | 0.104 | 0.748 |
| | Maximum | 0.315 | 0.740 | < 0.001 | 193 | 150167.7 | 3.229 | 0.074 |
| ii | All | 0.578 | 0.849 | < 0.001 | 197 | 8202.5 | 2.543 | 0.112 |
| | Truncated | 0.516 | 0.798 | < 0.001 | 197 | 11782.4 | 8.709 | **0.004** |
| | Maximum | 0.359 | 0.796 | < 0.001 | 199 | 198795.6 | 5.294 | **0.022** |
| iii | All | 0.729 | 0.751 | < 0.001 | 197 | 5991.0 | 1.019 | 0.314 |
| | Truncated | 0.696 | 0.693 | < 0.001 | 197 | 6326.0 | 1.900 | 0.170 |
| | Maximum | 0.404 | 0.665 | < 0.001 | 199 | 67330.0 | 3.976 | 0.048 |
| iv | All | 0.400 | 0.780 | < 0.001 | 193 | 11193.1 | 0.011 | 0.915 |
| | Truncated | 0.344 | 0.647 | < 0.001 | 169 | 15568.6 | 0.899 | 0.344 |
| | Maximum | 0.312 | 0.729 | < 0.001 | 199 | 198866.9 | 0.430 | 0.513 |
| v | All | 0.659 | 0.891 | < 0.001 | 199 | 7377.2 | 0.835 | 0.362 |
| | Truncated | 0.625 | 0.844 | < 0.001 | 197 | 8677.2 | 0.380 | 0.539 |
| | Maximum | 0.352 | 0.812 | < 0.001 | 197 | 182753.7 | 0.951 | 0.331 |

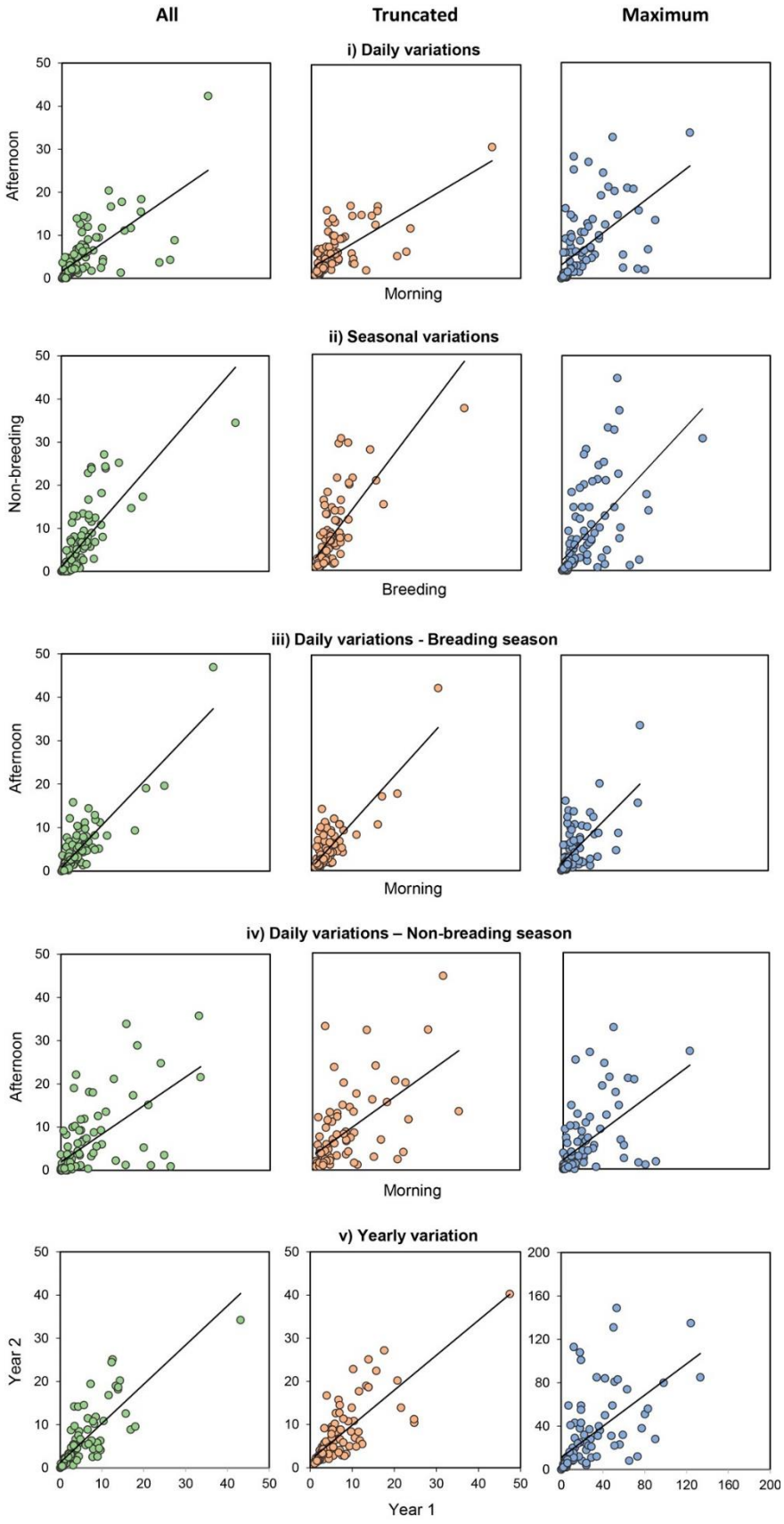



Figure 3.1 – Differences in the distribution of woodpigeon abundance (number of pigeons / 10 minutes of observation) on Terceira island, considering each dataset of abundance (All – all data, Truncated – excluding zero counts; Maximum – maximum abundance ever recorded), according to: i) daily variation; ii) seasonal variations, iii) daily variation during the breeding season; and iv) daily variation during the non-breeding season; v) yearly variation.

## 3.2. Mapping the distribution and abundance

Overall, the distribution pattern of the abundance observed during the Breeding season seems to be similar to the pattern found during the Non-breeding season (Figure 3.2). The standard deviation was frequently high, mainly for the higher values of observed abundance, sometimes surpassing those values.

During the Breeding season the higher values of abundance were confined to three main regions (northwest, southwest and east regions of the island), whereas during the Non-breeding season the higher values were more homogeneously distributed across the island.

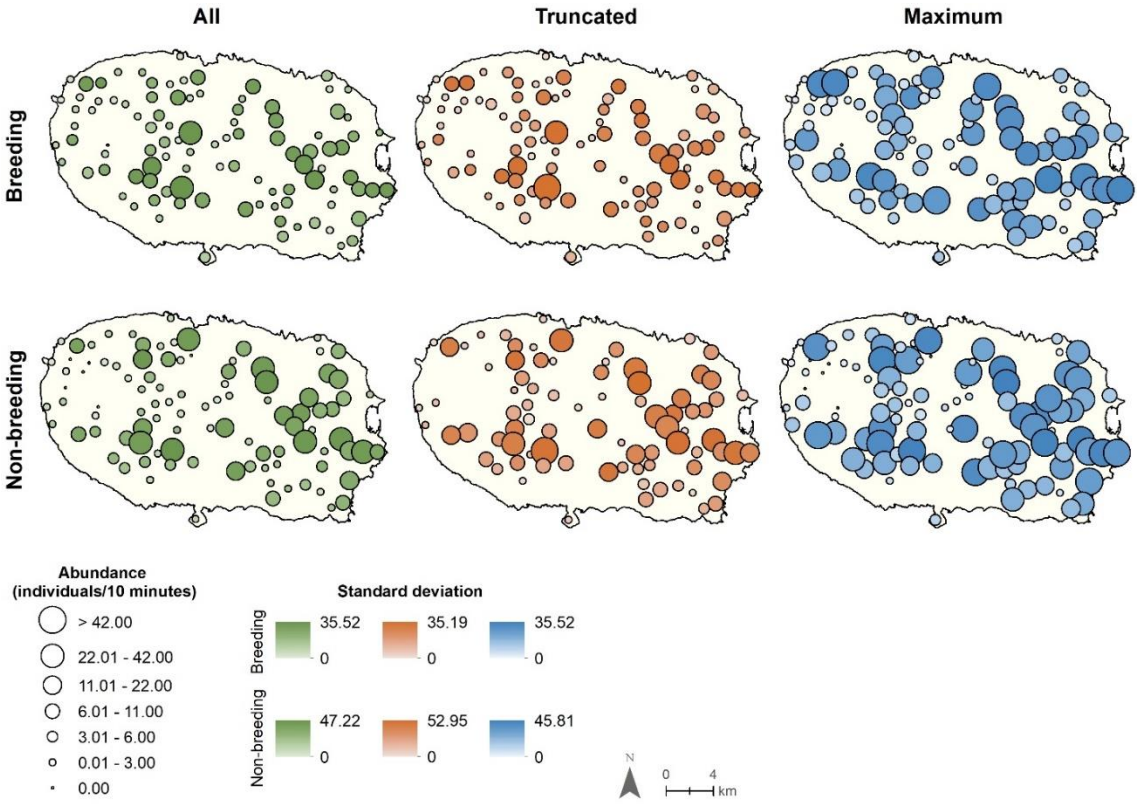



Figure 3.2 – Distribution of woodpigeon estimated abundance (number of pigeons / 10 minutes of observation) in breeding and non-breeding seasons, on Terceira island, considering each dataset of abundance (All - all data; Truncated - excluding zero counts; Maximum - maximum abundance ever recorded).

## 3.3. Modelling the distribution of abundance

All the transformations improved the correlations between observed and expected abundance relatively to the original data. From all the transformations we tested, the reciprocal transformation was the one that performed better for the Breeding season with *All data*, *Truncated data* and *Maximum data*. The $R^2$ values were 0.234, 0.278, and 0.266, respectively (Figure 3.3). For the Non-breeding season with *All data*, the best transformation was the Box-Cox transformation and the $R^2$ value was 0.333, while for the same season but with *Truncated data* and *Maximum data* the reciprocal transformation yields the best results again, with values of 0.379 and 0.338, respectively (Figure 3.3). Square root transformation was the test with lower $R^2$ values in most of the scenarios, with the exception for the Non-breeding season with *All data*, being the reciprocal transformation the test with lower $R^2$ (0.257 against 0.212) (Annexes 4, 5, 6, 7, 8, 9).

Regarding the final models, there was correlation between the observed abundance and the expected abundance in every scenario. However, the $R^2$ values of the Non-breeding season were higher than the correspondent values of the Breeding season (Table 6, Figure 3.3). The *Truncated data* exhibited the higher correlation between the observed and expected abundance, in both Breeding season and Non-breeding season. *Maximum data* and *All data* showed slightly lower correlation than the *Truncated data*, being the former the second best.

There was little variability between replicas, being the breeding season with *Truncated data* the one with higher variability (Figure 3.4).

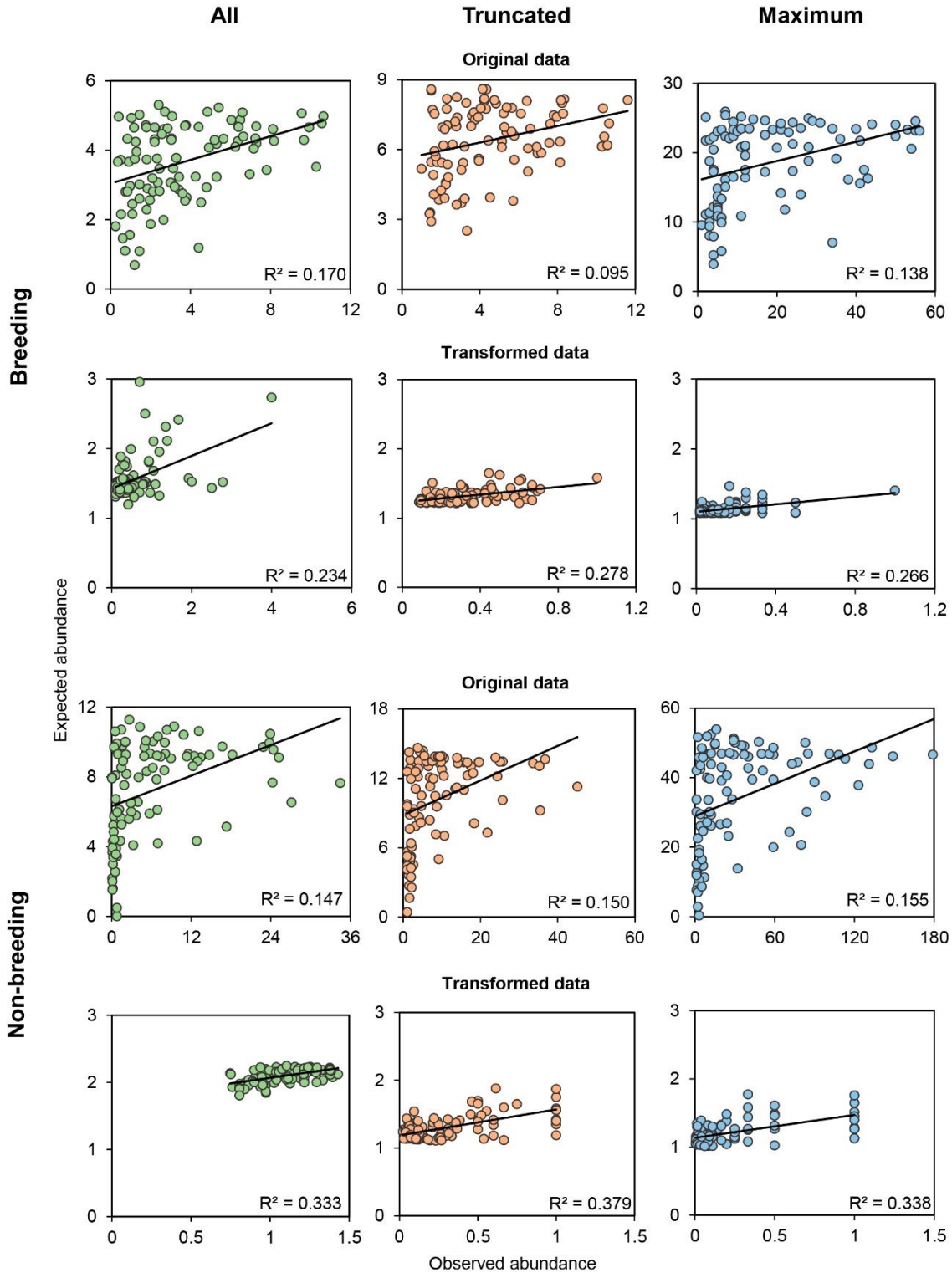



Figure 3.3 – Comparison between models performed for Breeding (top) and Non-breeding (bottom) seasons with the original data and a model performed with the best transformation of each season and abundance (number of pigeons / 10 minutes of observation) datasets (All - all data; Truncated - excluding zero counts; Maximum - maximum abundance ever recorded). All the transformations are "reciprocal transformations" ($1/x$), with the exception for the Non-breeding season with All data, which is the Box-Cox transformation $x(\boldsymbol{\lambda}) = \frac{x^{\lambda}-1}{\lambda}$ (λ = 0.101).

Table 6 – Results of the modelling for the Breeding season (B) and Non-breeding season (NB) and for the three types of abundance (number of pigeons / 10 minutes of observation) datasets (All - all data; Truncated - excluding zero counts; Maximum - maximum abundance ever recorded). The equations relating observed abundance with environmental factors, the R2 and the significance levels of each Principal Component (PC) for the relevant model, are given. All the models represented, were performed with the reciprocal transformation $(1/x)$, except for the NB All which was performed with the Box-Cox transformation $x(\lambda) = \frac{x^{\lambda}-1}{\lambda}$ (λ = 0.101). For the *Truncated data*, the dataset with better results, the variables the most contribute for each PC are the following: PC01 – Terrain ruggedness index and accumulated rainfall; PC02 – coniferous forests; PC03 – eucalyptus plantations and natural vegetation.

| Season | Abundance | Equation | $R^2$ | PC01 | PC02 | PC03 | PC04 | PC10 |
|---|---|---|---|---|---|---|---|---|
| B | All | 3.567 - 0.004*PC01 - 0.004*PC02 + 0.012*PC03 - 0.011*PC04 - 0.018*PC10 + 1 | 0.234 | < 0.001 | 0.058 | 0.003 | 0.036 | 0.105 |
| | Truncated | 0.601 - 0.001*PC01 - 0.003*PC02 + 0.003*PC03 + 1 | 0.278 | < 0.001 | < 0.001 | 0.003 | ------ | ------ |
| | Maximum | 0.443 - 0.001*PC01 - 0.002*PC02 + 0.003*PC03 - 0.002*PC04 + 1 | 0.265 | < 0.001 | < 0.001 | 0.003 | 0.070 | ------ |
| NB | All | 0.521 + 0.001*PC01 + 0.003*PC02 + 1 | 0.333 | < 0.001 | < 0.001 | ------ | ------ | ------ |
| | Truncated | 1.015 - 0.002*PC01 - 0.005*PC02 + 0.004*PC03 + 1 | 0.379 | < 0.001 | < 0.001 | 0.037 | ------ | ------ |
| | Maximum | 0.901 - 0.002*PC01 - 0.005*PC02 + 0.004*PC03 + 1 | 0.337 | < 0.001 | < 0.001 | 0.054 | ------ | ------ |

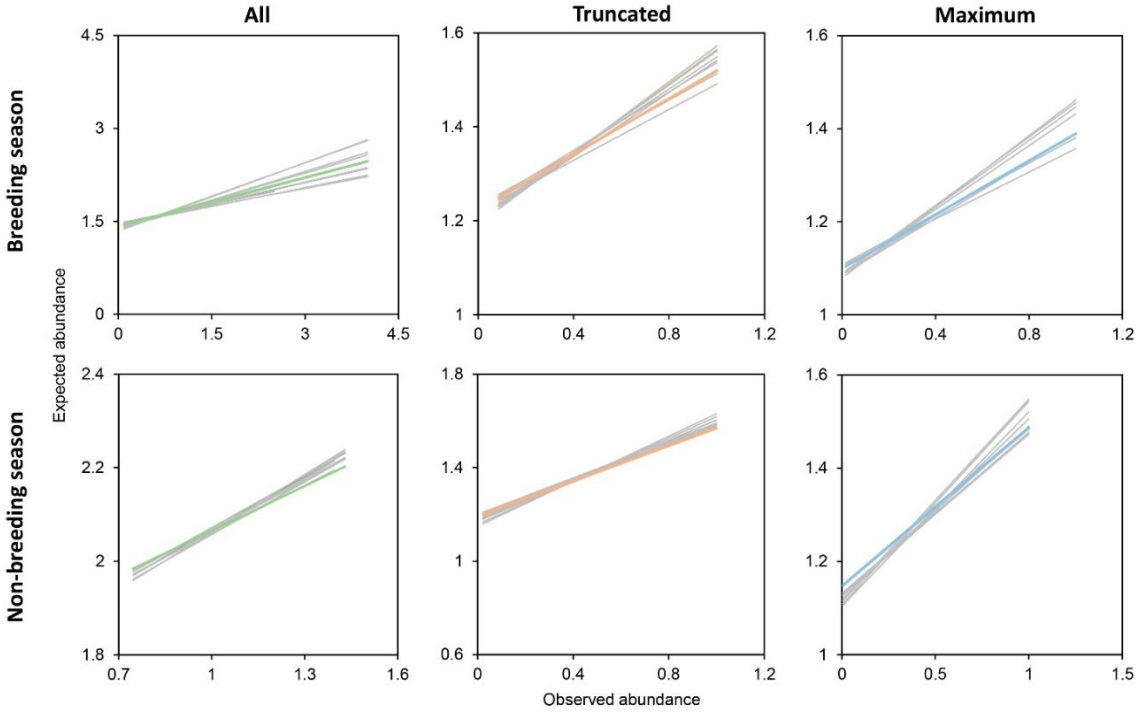



Figure 3.4 – Variation of each replica represented as trend lines of each dependent variable (number of pigeons / 10 minutes of observation) (All - all data; Truncated - excluding zero counts; Maximum - maximum abundance ever recorded). The coloured lines correspond to the trend lines of the final models. The y-axes and x-axes of the graphs start with other numbers than zero, for better visualization purposes.

The projection of the models based in the Breeding season indicates that the expected abundance levels do not match in many areas with the observed abundance for any of the abundance datasets (Figure 3.5). This relation seems to be improved in the Non-breeding season, as the $R^2$ values show (Table 6). The abundance dataset with higher correlation between the observed and expected abundance seems to be the *Truncated data* (Figure 3.5). Overall, the models seem to predict better the areas of low abundance in comparison to the areas of high abundance.

Focusing on the best results, achieved with the *Truncated data*, it was possible to verify that PC01, PC02 and PC03 were the components that contributed the most for the models, in both seasons (Table 6). The interpretation of the role of each component should take into account that the original data suffered a reciprocal transformation. Therefore, an increase in the values of the first two components will correspond to an increase in the values of the abundance (original data). For PC03 is the opposite. Previously (Table 4) we found that PC01 mainly represented a variation in TRI and accumulated rainfall (negative sign in both cases), and PC02 a variation in coniferous forests (mainly pine; positive sign); PC03 represented a variation on natural vegetation and eucalyptus plantations, but with opposite signs: positive and negative, respectively.

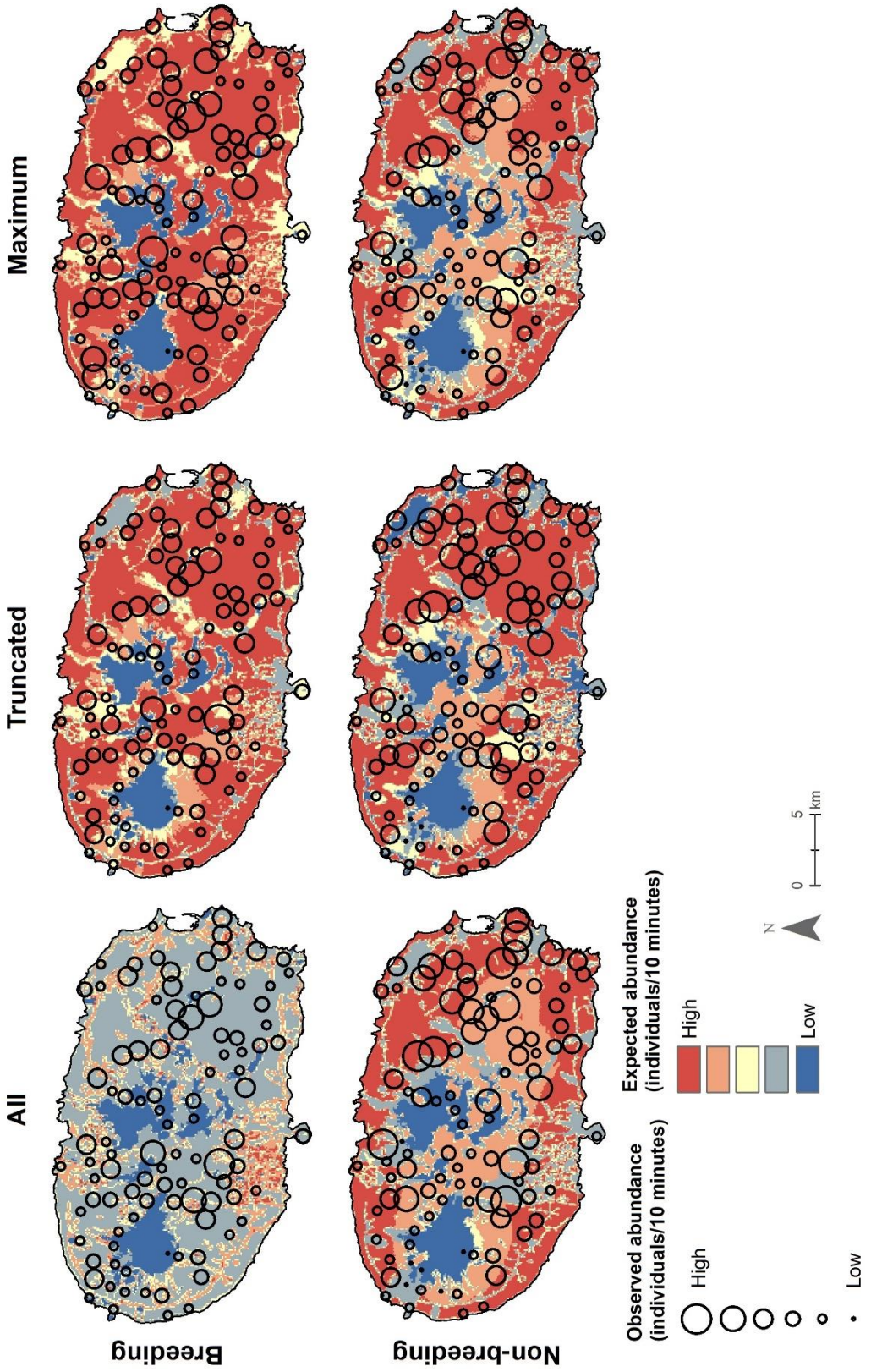



Figure 3.5 – Distribution of observed and expected woodpigeon abundance (number of pigeons / 10 minutes of observation) in Terceira island. Expected abundance based on regression models relating observed abundance with environmental factors. The expected and observed abundance (All - all data; Truncated - excluding zero counts; Maximum - maximum abundance ever recorded) are represented without the transformations.

Comparing the Breeding season and the Non-breeding season models performed with the *Truncated data* with an adjusted and comparable scale (Figure 3.6), one can assume that the expected abundance distribution pattern is quite similar between the two, but the Breeding season model shows lower levels of abundance overall. The regions with lower values of expected abundance were the regions with natural vegetation, with high TRI values (most likely there is a strong correlation between these two) and urban and social areas, while the agriculture fields had higher values of expected abundance (Figure 3.6, ANNEX 2).

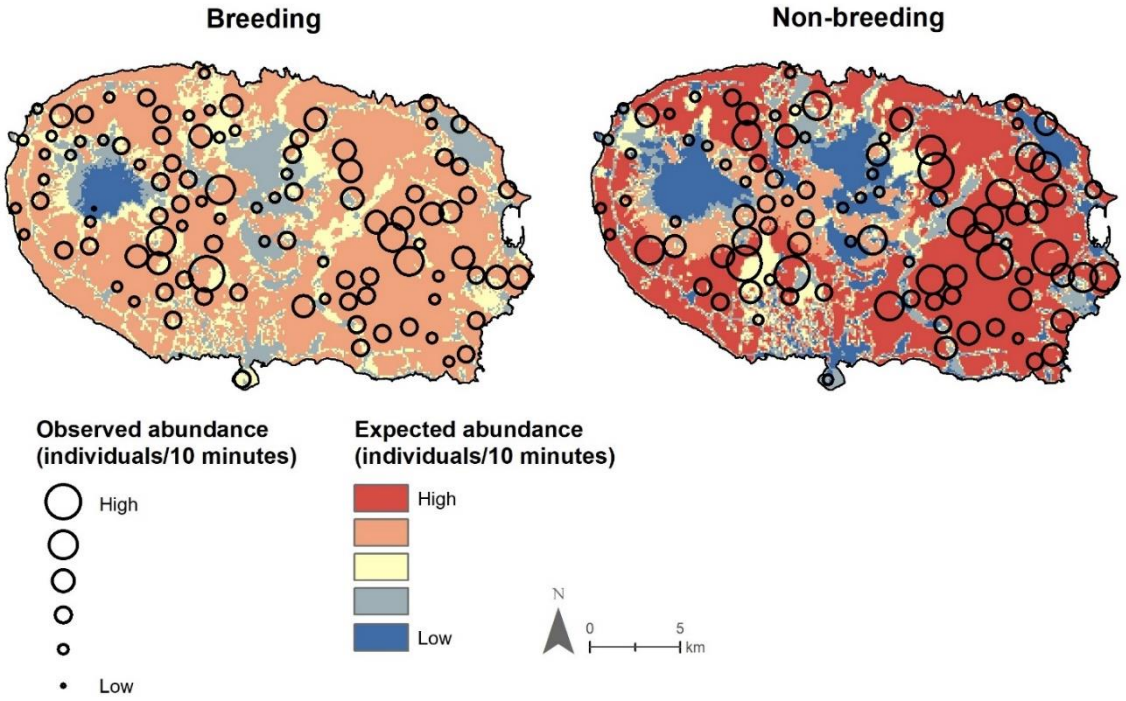



Figure 3.6 – Distribution of observed and expected woodpigeon abundance (number of pigeons / 10 minutes of observation) (*Truncated data* – excluding zero counts) on Terceira island. Expected abundance based on regression models relating observed abundance with environmental factors. The expected and observed abundance are represented with the respective transformations.

# 4. Discussion.

With this work we explored ecological aspects never studied before in the Azores woodpigeon, more specifically the habitat selection across Terceira island during the Breeding season and the Non-breeding season. We also compared three different abundance estimates.

We will discuss, in first place all the results about the abundance datasets used and the uncertainty inherent to them and the transformations made for the modelling process. After, we discuss the subjects related to the habitat selection of the woodpigeon and the prediction of the distribution of abundance through the island.

## 4.1. Abundance estimations

The abundance estimations based in *All data* and *Truncated data*, in comparison to *Maximum data*, were the ones that exhibited the highest $R^2$ and Spearman rank correlation coefficients in the tests of daily, seasonal, and yearly variations in abundance. Regarding the models of the distribution of abundance, the higher correlation between observed abundance and expected abundance was observed with the *Truncated data*. Pearce and Ferrier (2001) developed distinct regression models for many different species and all of them were performed with two different data: including all sites independently of the values of abundance and including only the sites with abundance higher than zero. Although Pearce and Ferrier (2001) found that the models perform better when all values of abundance are taking into account, we conclude that the *Truncated* abundance performed better than the *All data* and *Maximum data* abundance. The reason for this difference in results can be due to the high frequency of zeros in our dataset: in 35% of the sampling counts no individuals were registered. Such high prevalence of null observations may create biases, which will cause uncertainties in parameter estimates and in statistical inferences (Martin et al., 2005).

The observed abundance was associated with high values of standard deviation. In many of the sampling points, there were substantial variations in the recorded number of pigeons (e.g. the observed abundance of point ranges from 0 to 179, just in the non-breeding season). This pattern was translated in standard deviation values being very similar to, or even higher (especially in the points with high abundances), than the values of the respective observed abundance. In addition, we did not consider the species detectability, which may influence the recorded abundance and can cause

uncertainties in its estimation. Detectability may vary with the observer characteristics, habitat type, environmental conditions affecting the level of fauna activity, season and time of the day (Pearce and Ferrier, 2001). Although it is impossible to eliminate all uncertainties associated to detectability, most of the variation sources were as controlled as possible: i) the observer was always the same person; ii) the sampling point counts were always performed in the exact same place; iii) counts were always performed with good visibility conditions; and iv) the sampling time of the day was always respected. In addition, the detectability increases with the abundance of a species (McCarthy et al., 2013). Given that the Azores woodpigeon is relatively abundant on Terceira and the procedures followed in the sampling protocol, we assume that detectability likely did not significantly affect the estimations of abundance.

All tested transformations of abundance improved the $R^2$ values when compared to the original data. Reciprocal transformation was the best data transformation for all the cases, except for the abundance with *All data* for the Non-breeding season, which was the Box-Cox transformation.

Acknowledging these results, the following discussion is based on the results obtained with the *Truncated data*.

## 4.2. Annual and daily variation on abundance

The results of the analyses of variance suggested that there were no significant differences in observed abundance on Terceira between years and between parts of the day. In a study about the feeding habits of the woodpigeon (Murton et al., 1964), in Carlton (United Kingdom), between January and March, birds were observed feeding in the fields all day, and during the breeding season they spent more time resting in the woods, because they required a shorter feeding day. Although no statistical tests were done comparing relative abundance of woodpigeon between different parts of the day (i.e. morning versus afternoon), their work suggests changes in birds' abundance through the space during the day, in the breeding season and no differences during winter. These results contrast partially with those for the Azorean woodpigeons, where daily abundance differences were not found in any of the two seasons tested. Lucas (2004), in a work about the ecology of Azores woodpigeon on Terceira, also did not observe differences in the distribution of abundance between periods of the day during winter. Differences in abundance between parts of the day were expected, as it happens with other bird species (King and Dallimer, 2003). Usually, birds tend to be

more active during the morning than during the afternoon, but the results of the Azores woodpigeon point to a homogeneous abundance distribution throughout the area of study between those parts of the day. We did not expect differences between the first and second year, because in preliminary tests (not shown) performed with data from one meteorological station from Terceira, no difference was observed on temperatures and rainfall between the two years.

## 4.3. Seasonal patterns in abundance distribution

During the Breeding season the higher values of observed abundance are concentrated in three main regions: northwest, southwest and east regions of the island. However, during the Non-breeding season, the higher abundance values are more distributed across the island. As described before, the woodpigeon in the beginning of autumn may join in feeding grounds (Murton, 1965). During the breeding season they spend more time in forested areas (Murton et al., 1964) and that can explain why during this season the distribution of Azores woodpigeon abundance forms spatial clusters.

Regarding the models of abundance distribution, the Non-breeding season model was better adjusted to the observed data (highest $R^2$) than the Breeding season model. Both models selected the same principle components (PC01, PC02, PC03) as the ones with highest explanatory power, which suggests that differences in the ecological niches between seasons should be minimal. The explanatory variables most related to the Azores woodpigeon abundance were, positively, the percentage of coniferous forests and eucalyptus plantations, negatively, the percentage of natural vegetation, the accumulated rainfall in the months from September to April and the Terrain ruggedness index (TRI).

The positive relationship of percentage of eucalyptus with the abundance of the Azores woodpigeon contrasts with the situation of other bird species in continental Portugal, where the diversity and abundance of birds is lower in eucalyptus plantations than in other habitats (Proença et al., 2010). According to Equipa Atlas (2008), the Azores woodpigeon inhabits coniferous forests, which is in agreement with our results, once the percentage of coniferous forests has a positive effect on abundance. The reason for the observed positive correlations may rely on the fact that the species needs tall trees to build the nests and to roost (Murton et al., 1964; Murton, 1965). TRI exhibited a negative relationship with the abundance of the Azores woodpigeon and the same

pattern was observed for the accumulated rainfall in the months from September to April and percentage of natural vegetation. Rainfall is usually correlated with altitude (Brunsdon et al., 2001), and the latter should be partially correlated with TRI (the higher values of TRI correspond to the higher areas of the island). The negative relation between the percentage of natural vegetation and the abundance of the Azores woodpigeon was not expected. Probably, because on Terceira island, the areas with natural vegetation are mainly confined to localities with high TRI and accumulated rainfall (Silva and Smith, 2006; Elias et al., 2016).

These results suggest an altitudinal zonation of the Azores woodpigeon abundance, with areas with lower average altitude levels presenting higher values of abundance, as stated before by Lucas (2004). Further studies should be done to address this question. Pico island would be a good study area because there, natural vegetation is not restricted to higher altitudes (IFAA, 2007).

Agriculture fields had low contribution to any of the first 10 PCs, which were the ones most related with the environmental variability of Terceira island and that were used to build regression models. Still, regression models (Figure 3.7) suggest high expected abundance in areas that correspond to agriculture fields in both seasons. Lucas (2004) also found that during the winter the Azores woodpigeon is more abundant in agriculture fields on Terceira, which should be related to the dietary habits of the species. In continental Europe, the woodpigeon gathers in large numbers in feeding areas during the non-breeding season, whereas during the breeding season spend more time closer to the nests (Murton and Isaacson, 1962; Murton et al., 1964). As expected, the same pattern was observed on Terceira, i.e. in forested areas there are higher abundances during the Breeding season in comparison to the Non-breeding season; in open areas (agriculture fields) higher abundances are observed during the Non-breeding season than during the Breeding season. Unexpectedly, the abundance was lower in natural vegetation areas, but that might be due to the reasons discussed above. It was also low in social and urban areas. Although the woodpigeon is an ecologically dynamic and outstandingly adaptable species, it needs habitats that provide shelter and food (Cramp, 1985) and those areas where the Azores woodpigeon is less abundant might not provide those conditions.

We consider that the models performed well, however we think they could be further improved with other variables that might be important to explain the Azores woodpigeon distribution in Azores, as the food distribution across space and normalized difference vegetation index. Also, two PCAs could be done, one for each

season, because the environmental variables can have some differences between those periods that were not depicted with a single PCA.

## 4.4 Final remarks

Our results suggest that the Azores woodpigeon is quite abundant in agriculture fields in both seasons. Additional studies on woodpigeon ecology (e.g. feeding habits) are needed to evaluate the importance of the species on eventual damages in crops reported by local farmers. An additional study should be done to assess whether or not altitude explains the negative relation between the abundance of the Azores woodpigeon and the percentage of natural vegetation.

This is one of the first studies of this nature for this species, not only in Azores but also in continental Europe. We see potential in using abundance as state variable to build ecological models, mostly for species with high abundances and large distribution ranges, but it is important to include the variables that explain better that distribution.

# 5. References

Adler, G.H., Austin, C.C., and Dudley, R. (1995). Dispersal and speciation of skinks among archipelagos in the tropical Pacific Ocean. Evol. Ecol. 9, 529–541.

Anderson, R.P., Peterson, A.T., and Gómez-Laverde, M. (2002). Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. Oikos 98, 3–16.

Andrade, P., Rodrigues, P., Lopes, R.J., Ramos, J.A., Cunha, R.T., and Gonçalves, D. (2015). Ecomorphological patterns in the Blackcap *Sylvia atricapilla*: insular versus mainland populations. Bird Study 62,498–507.

Andrade, P.A.M., Rodrigues, T.M., Fontaine, R., Neves, V.R.C., Fonseca, A., Alves, P.C., Carneiro, M.J.P., and Gonçalves, D. (2016). Genetic and morphological differentiation in an island bird, the Azorean Woodpigeon (*Columba palumbus azorica*). pp. 365–366 *in*: Gabriel, R., Elias, R.B., Amorim, I.R., and Borges, P.A.V. (Eds). Conference program and abstracts of the 2nd International Conference on Island Evolution, Ecology and Conservation: Island Biology 2016, 18-22 July 2016, Angra do Heroísmo, Azores, Portugal. Arquipélago. Life and Marine Sciences. Supplement 9.

Andrews, D.F. (1971). A note on the selection of data transformations. Biometrika 58, 249–254.

Atkinson, A.C. (1973). Testing transformations to normality. J. R. Stat. Soc. Series B (Methodological) 35, 473–479.

Azevedo, E.B. (1996). Modelação do Clima Insular à Escala Local. Modelo CIELO aplicado à ilha Terceira. PhD Thesis, University of Azores. [in Portuguese]

Bannerman, D.A., and Bannerman, W.M. (1966). Birds of the Atlantic Islands. Vol. III - A History of the Birds of the Azores. (Edinburgh: Oliver and Boyd).

Barton, K. (2015). MuMIn: multi-model inference. R package version 1.15.6.

Bea, A., Beitia, R., and Fernández, J.M. (2003). The census and distribution of wintering woodpigeons *Columba palumbus* in the Iberian Peninsula. Ornis Hungarica 12–13, 157–167.

BirdLife International (2016). *Columba palumbus*. The IUCN Red List of Threatened Species 2016: e.T22690103A86077750.

BirdLife International and Handbook of the Birds of the World (2016). *Columba palumbus*. The IUCN Red List of Threatened Species. Version 2017-1.

Borges, P.A.V., and Brown, V.K. (1999). Effect of island geological age on the arthropod species richness of Azorean pastures. Biol. J. Linn. Soc. 66, 373–410.

Box, G.E.P., and Cox, D.R. (1964). An analysis of transformations. J. R. Stat. Soc. Series B (Methodological) 26, 211–252.

Brunsdon, C., McClatchey, J., and Unwin, D.J. (2001). Spatial variations in the average rainfall–altitude relationship in Great Britain: an approach using geographically weighted regression. Int. J. Clim. 21, 455–466.

Cabral, M.J., Almeida, J., Almeida, P.R., Dellinger, T., Almeida, N.F., Oliveira, M.E., Palmeirim, J.M., Queirós, A.I., Rogado, L., and Santos-Reis, M. (2005). Livro Vermelho dos Vertebrados de Portugal (Lisboa, Portugal: Instituto da Conservação da Natureza). [in Portuguese]

Carroll, R.J., and Ruppert, D. (1981). On prediction and the power transformation family. Biometrika 3, 609–615.

Carroll, S.S., and Pearson, D.L. (1998). The effects of scale and sample size on the accuracy of spatial predictions of Tiger Beetle (Cicindelidae) species richness. Ecography (Cop.). 21, 401–414.

Cataldo, D. (2017) Phenotypic differentiation in the Azorean woodpigeon (Columba palumbus azorica). MSc Thesis in Ecology and Environment, University of Porto.

Catry, P., Costa, H., Elias, G., and Matias, R. (2010). Aves de Portugal. Ornitologia do Território Continental (Lisboa: Assírio & Alvim). [in Portuguese]

Connor, S.E., van Leeuwen, J.F.N., Rittenour, T.M., van der Knaap, W.O., Ammann, B., and Björck, S. (2012). The ecological impact of oceanic island colonization – a palaeoecological perspective from the Azores. J. Biogeogr. 39, 1007–1023.

Cox, C.B., Moore, P.D., and Ladle, R. (2005). Biogeography: An Ecological and Evolutionary Approach (Oxford: Blackwell Publishing).

Cramp, S. (1985). Handbook of the birds of Europe the Middle East and North Africa, the birds of the Western Palearctic: Terns to Woodpeckers. Volume IV. (Oxford: Oxford University Press).

Dias, E., and Melo, C. (2010). Factors influencing the distribution of Azorean mountain vegetation: implications for nature conservation. Biodivers. Conserv. 19, 3311–3326.

Dickens, M., and Neves, V. (2005). Post-breeding density and habitat preferences of the Azores woodpigeon, *Columba palumbus azorica*: an inter-island comparison. Arquipélago Life and Marine Sciences 22A, 61–69.

Dourado, C.G., Duarte, M.A., Grosso, A.R., Bastos-Silveira, C., Marrero, P., Oliveira, P., Paulo, O.S., and Dias, D. (2013). Phylogenetic origin of the endemic pigeons

from Madeira (*Columba trocaz*) and Azores Islands (*Columba palumbus azorica*). J. Ornithol. 155, 71–82.

Elias, R.B., Gil, A., Silva, L., Fernández-Palacios, J.M., Azevedo, E.B., and Reis, F. (2016). Natural zonal vegetation of the Azores Islands: characterization and potential distribution. Phytocoenologia 46, 107–123.

Elith, J., and Leathwick, J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697.

Equipa Atlas (2008). Atlas das Aves Nidificantes em Portugal (1999-2005). (Lisboa, Portugal: Assírio & Alvim).

Fontaine, R., Rodrigues, T., Neves, V., Fonseca, A., and Gonçalves, D. (2014). Ecologia do pombo-torcaz dos Açores *Columba palumbus azorica* – Resultados de um ano de monitorização na ilha Terceira. In: SPEA 2014. Livro de resumos do VIII Congresso de Ornitologia da SPEA – 1ª Edição. Sociedade Portuguesa para o Estudo das Aves. pp. 83–84. [in Portuguese]

Fontaine, R., Rodrigues, T.M., Fonseca, A., Neves, V., and Gonçalves, D. (2014). Relatório do 1º Censo de pombo-torcaz *Columba palumbus azorica* (Hartert, 1905) realizado em todas as ilhas do arquipélago dos Açores (outubro/novembro de 2014). Projeto AZORPI (M2.1.2/I/025/2011) CIBIO/InBIO University of Porto, University of Azores. [in Portuguese]

Francisco-Ortega, J., Santos-Guerra, A., Kim, S.C., and Crawford, D.J. (2000). Plant genetic diversity in the Canary Islands: a conservation perspective. Am J Bot 87, 909–919.

Franklin, J. (2010). Mapping Species Distributions. Spatial Inference and Prediction. (Cambridge: Cambridge University Press).

Fraser, D.A.S. (1967). Data transformations and the linear model. Ann. Math. Stat. 38, 1456–1465.

Gillespie, R.G., and Clague, D.A. (2009). Encyclopedia of Islands (London: University of California Press, Ltd).

Guisan, A., and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8, 993–1009.

Harrell, F.E. (2017). Hmisc: Harrell miscellaneous. R package version 4.0-3.

Hartert, E., and Ogilvie-Grant, W.R. (1905). On the birds of the Azores. Novitates Zoologicae 12, 80–128.

Haynes, P.J., Inglis, I.R., Isaacson, T.J., and Fryday, S.L. (2003). Woodpigeon *Columba palumbus* movements in eastern England. Bird Study 50, 33–38.

Hernandez, P.A., Graham, C.H., Master, L.L., and Albert, D.L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29, 773–785.

Howard, C., Stephens, P.A., Pearce-Higgins, J.W., Gregory, R.D., and Willis, S.G. (2014). Improving species distribution models: the value of data on abundance. Methods Ecol. Evol. 5, 506-513.

IFRAA (2007). Inventário Florestal da Região Autónoma dos Açores – 2007. Accessed in: 20/09/2016, at: http://drrf-sraa.azores.gov.pt/areas/inventario-florestal/Paginas/Download.aspx

Inglis, I.R., Isaacson, A.J., and Thearle, R.J.P. (1994). Long term changes in the breeding biology of the woodpigeon *Columba palumbus* in eastern England. Ecography 17, 182–188.

Johnson, T.H., and Stattersfield, A.J. (1990). A global review of island endemic birds. Ibis 132, 167–180.

Juan, C., Emerson, B.C., Oromí, P., and Hewitt, G.M. (2000). Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends Ecol. Evol. 15, 104–109.

Kéry, M., Royle, J.A., and Schmid, H. (2005). Modeling avian abundance from replicated counts using binomial mixture models. Ecol. Appl. 15, 1450–1461.

Kier, G., Kreft, H., Lee, T.M., Jetz, W., Ibisch, P.L., Nowicki, C., Mutke, J., and Barthlott, W. (2009). A global assessment of endemism and species richness across island and mainland regions. PNAS 106, 9322–9327.

King, T., and Dallimer, M. (2003). Daily activity, moult and morphometrics of the birds of São Tomé and Príncipe. Bull. African Bird Club 10, 84–93.

Leathwick, J.R., and Austin, M.P. (2001). Competitive interactions between tree species in New Zealand's old-growth indigenous forests. Ecology 82, 2560–2573.

Lee, D.C., and Marsden, S.J. (2008). Adjusting count period strategies to improve the accuracy of forest bird abundance estimates from point transect distance sampling surveys. Ibis 150, 315–325.

Lee, P.C., and Priston, N.E.C. (2005). Human attitudes to primates: perceptions of pests, conflict and consequences for primate conservation. pp 1–23 *in*: Paterson, J.D., and Wallis, J. (Eds). Primate-Human Interaction and Conservation (Alberta: American Society of Primatologists Publications).

Lobo, J.M., Lumaret, J.-P., and Jay-Robert, P. (2002). Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the

predictive capacity of different groups of explanatory variables. Glob. Ecol. Biogeogr. 11, 265–277.

Lomolino, M.V. (2005). Body size evolution in insular vertebrates: generality of the island rule. J. Biogeogr. 32, 1683–1699.

Lucas, L.M.N. (2004). Elementos da Ecologia do Pombo-torcaz dos Açores - *Columba palumbus torcaz*, Hartert 1905 - como Contributo para a Gestão e Conservação das suas Populações. Internship report. University of Lisbon. [in Portuguese]

Lyashevska, O., Brus, D.J., and van der Meer, J. (2016). Mapping species abundance by a spatial zero-inflated Poisson model: A case study in the Wadden Sea, the Netherlands. Ecol. Evol. 6, 532–543.

Marrero, P., Oliveira, P., and Nogales, M. (2004). Diet of the endemic Madeira Laurel Pigeon *Columba trocaz* in agricultural and forest areas: implications for conservation. Bird Conserv. Int. 14, 165–172.

Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J., and Possingham, H.P. (2005). Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecol. Lett. 8, 1235–1246.

McCarthy, M.A., Moore, J.L., Morris, W.K., Parris, K.M., Garrard, G.E., Vesk, P.A., Rumpff, L., Giljohann, K.M., Camac, J.S., Bau, S.S., Friend, T., Harrison, B., and Yue, B. (2013). The influence of abundance on detectability. Oikos 122, 717–726.

Molles, M.C. (2013). Ecology: Concepts and Applications (New York: Mc Graw-Hill International Edition).

Murton, R.K. (1965). The Wood-Pigeon. New Naturalist Monograph No 20. (London: Collins).

Murton, R.K., and Isaacson, A.J. (1962). The functional basis of some behaviour in the wood pigeon *Columba palumbus*. Ibis. 104, 503–521.

Murton, R.K., Westwood, N.J., and Isaacson, A.J. (1964). The feeding habits of the woodpigeon *Columba palumbus*, stock dove *C. oenas* and turtle dove *Streptopelia turtur*. Ibis. 106, 174–188.

Nally, R. Mac (2000). Regression and model-building in conservation biology, biogeography and ecology: the distinction between – and reconciliation of – 'predictive'and 'explanatory' models. Biodivers. Conserv. 9, 655–671.

Oliveira, P., Jones, M., Caires, D., and Menezes, D. (1999). Population trends and status of the Madeira Laurel Pigeon *Columba trocaz*. Bird Conserv. Int. 9, 387–395.

Oliveira, P., Menezes, D., Jones, M., and Nogales, M. (2006). The influence of fruit abundance on the use of forest and cultivated field habitats by the endemic Madeira laurel pigeon *Columba trocaz*: implications for conservation. Biol. Cons. 130, 538–548.

Paulay, G. (1994). Biodiversity on oceanic islands: its origin and extinction. Am. Zool. 34, 134–144.

Pearce, J., and Ferrier, S. (2001). The practical value of modeling relative abundance of species for regional conservation planning: a case study. Biol. Conserv. 98, 33–43.

Proença, V.M., Pereira, H.M., Guilherme, J., and Vicente, L. (2010). Plant and bird diversity in natural forests and in native and exotic plantations in NW Portugal. Acta Oecologica 36, 219–226.

Purroy, J., and Purroy, F. J. (2016). Paloma torcaz – *Columba palumbus*. *in*: Salvador, A., and Morales, M. B. (Eds.). Enciclopedia Virtual de los Vertebrados Españoles. (Madrid: Museo Nacional de Ciencias Naturales). Accessed in: 28/09/2017, at: http://www.vertebradosibericos.org/ [in Spanish]

Rando, J.C., Alcover, J.A., Olson, S.L., and Pieper, H. (2013). A new species of extinct scops owl (Aves: Strigiformes: Strigidae: *Otus*) from São Miguel Island (Azores Archipelago, North Atlantic Ocean). Zootaxa 3647, 343–357.

Rovero, F., Martin, E., Rosa, M., Ahumada, J.A., and Spitale, D. (2014). Estimating species richness and modelling habitat preferences of tropical forest mammals from camera trap data. PLoS One 9, e110971.

Rull, V., Lara, A., Rubio-Inglés, M.J., Giralt, S., Gonçalves, V., Raposeiro, P., Hernández, A., Sánchez-López, G., Vázquez-Loureiro, D., Bao, R., Masqyé, P., and Sáez, A. (2017). Vegetation and landscape dynamics under natural and anthropogenic forcing on the Azores Islands: a 700-year pollen record from the São Miguel island. Quat. Sci. Rev. 159, 155–168.

Silva, L., and Smith, C.W. (2006). A quantitative approach to the study of non-indigenous plants: an example from the Azores archipelago. Biodivers. Conserv. 15, 1661–1679.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Tognelli, M.F., and Kelt, D.A. (2004). Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. Ecography 27, 427–436.

van Valen, L. (1973). Pattern and the balance of nature. Evol Theory 1, 31–49.

Vaurie, C. (1961). Systematic notes on Palearctic birds Nr. 48, Columbidae: The genus *Columba*. Amer. Mus. Nov. 2043, 1–12.

Whittaker, R., and Fernández-Palacios, J.M. (2007). Island Biogeography. Ecology, Evolution and Conservation. (New York: Oxford University Press).

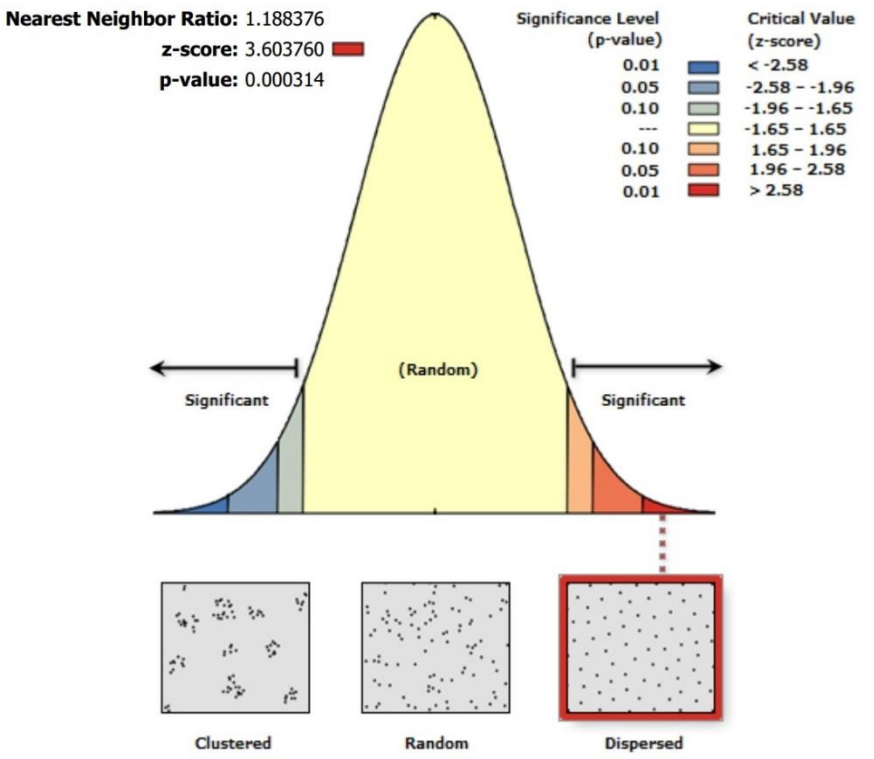
Wiens, J.A. (2002). Predicting species occurrences: progress, problems, and prospects. pp. 739–749 *in:* Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., and Samson, F.B. (Eds.). Predicting Species Occurrences: Issues of Accuracy and Scale. (Washington: Island Press).

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., NCEAS Predicting Species Distributions Working Group. (2008). Effects of sample size on the performance of species distribution models. Divers. Distrib. 14, 763–773.

# ANNEXES

ANNEX 1 – The nearest neighbour index measures the clustering level of the sampling points. The sampling points were classified by this test as dispersed.

## Average Nearest Neighbor Summary

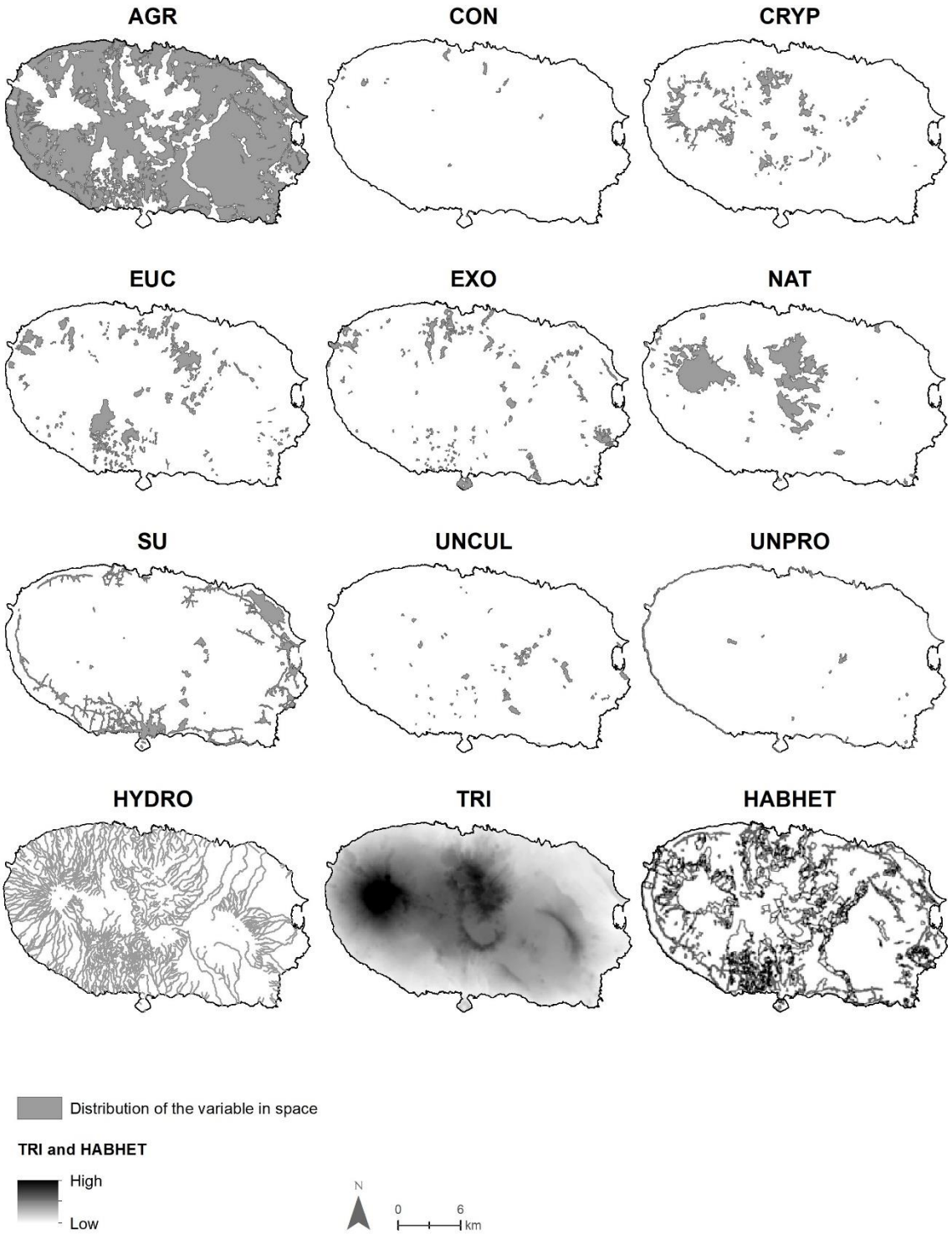



Given the z-score of 3.60, there is a less than 1% likelihood that this dispersed pattern could be the result of random chance.

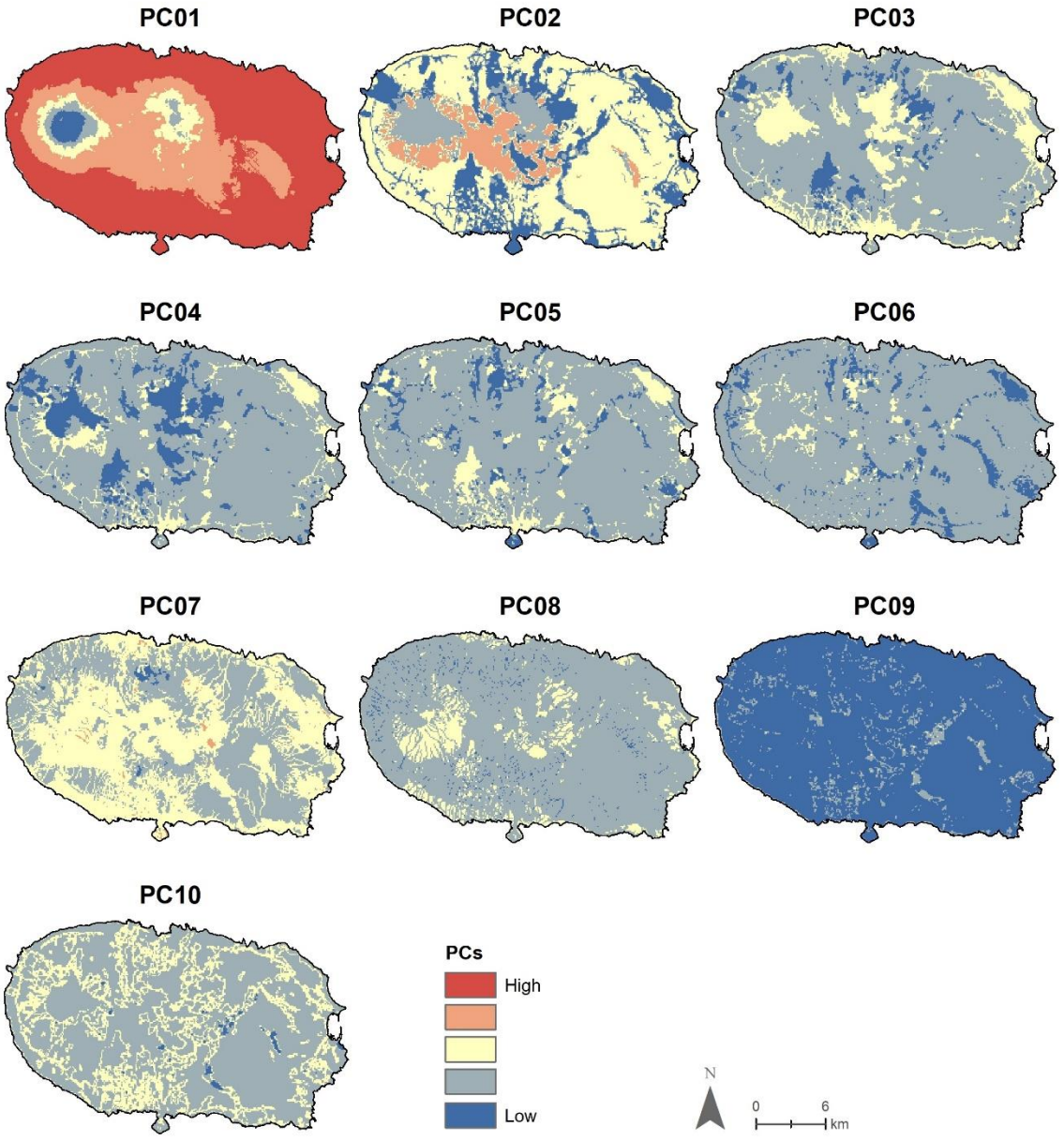
## Average Nearest Neighbor Summary

| | |
|---|---|
| **Observed Mean Distance:** | 1288.931888 Meters |
| **Expected Mean Distance:** | 1084.616447 Meters |
| **Nearest Neighbor Ratio:** | 1.188376 |
| **z-score:** | 3.603760 |
| **p-value:** | 0.000314 |

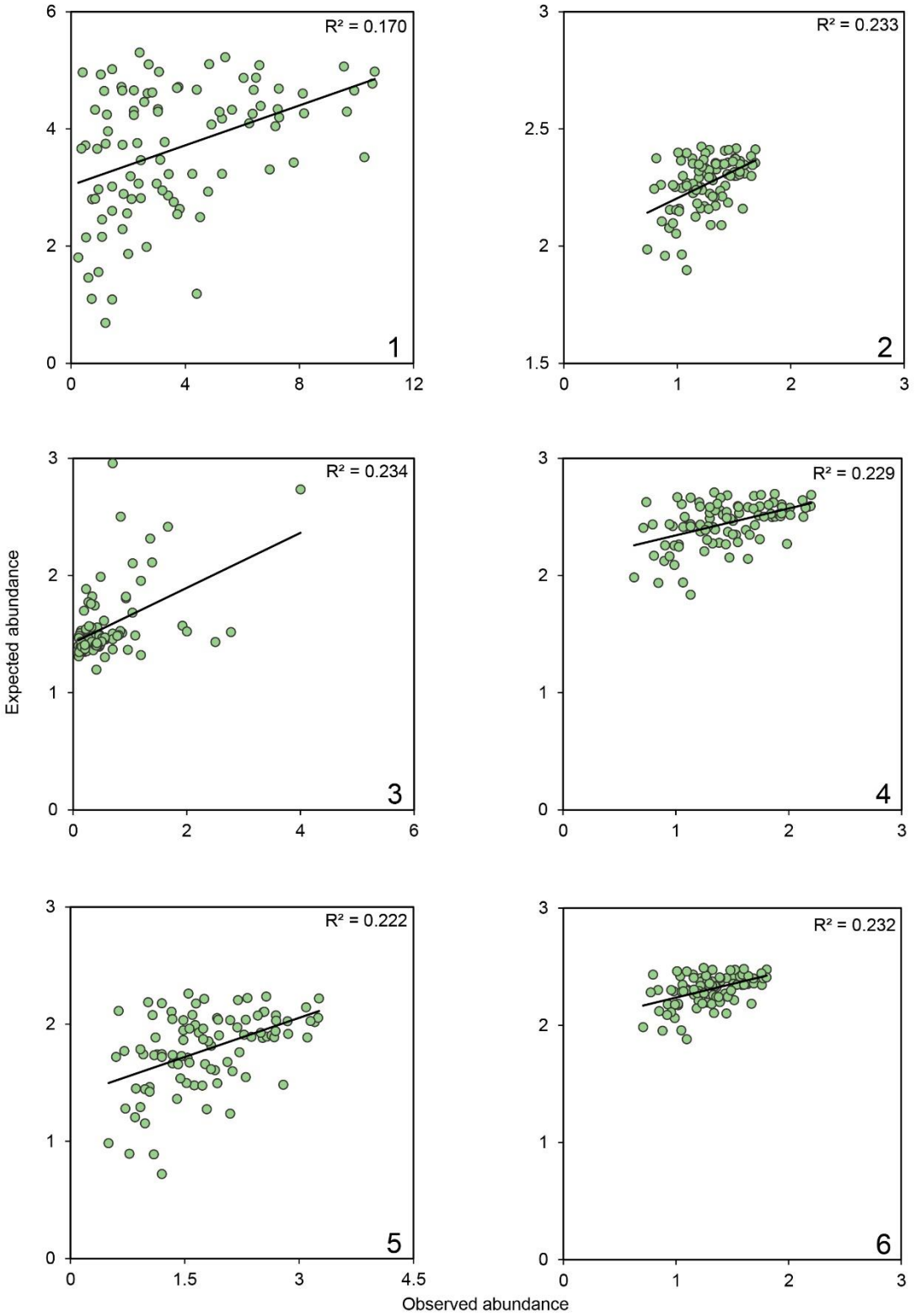
ANNEX 2 – Spatial representation of habitat variables (grey), hydrographic network and Terrain ruggedness index on Terceira (Azores). Habitat categories are explained in Table 2.



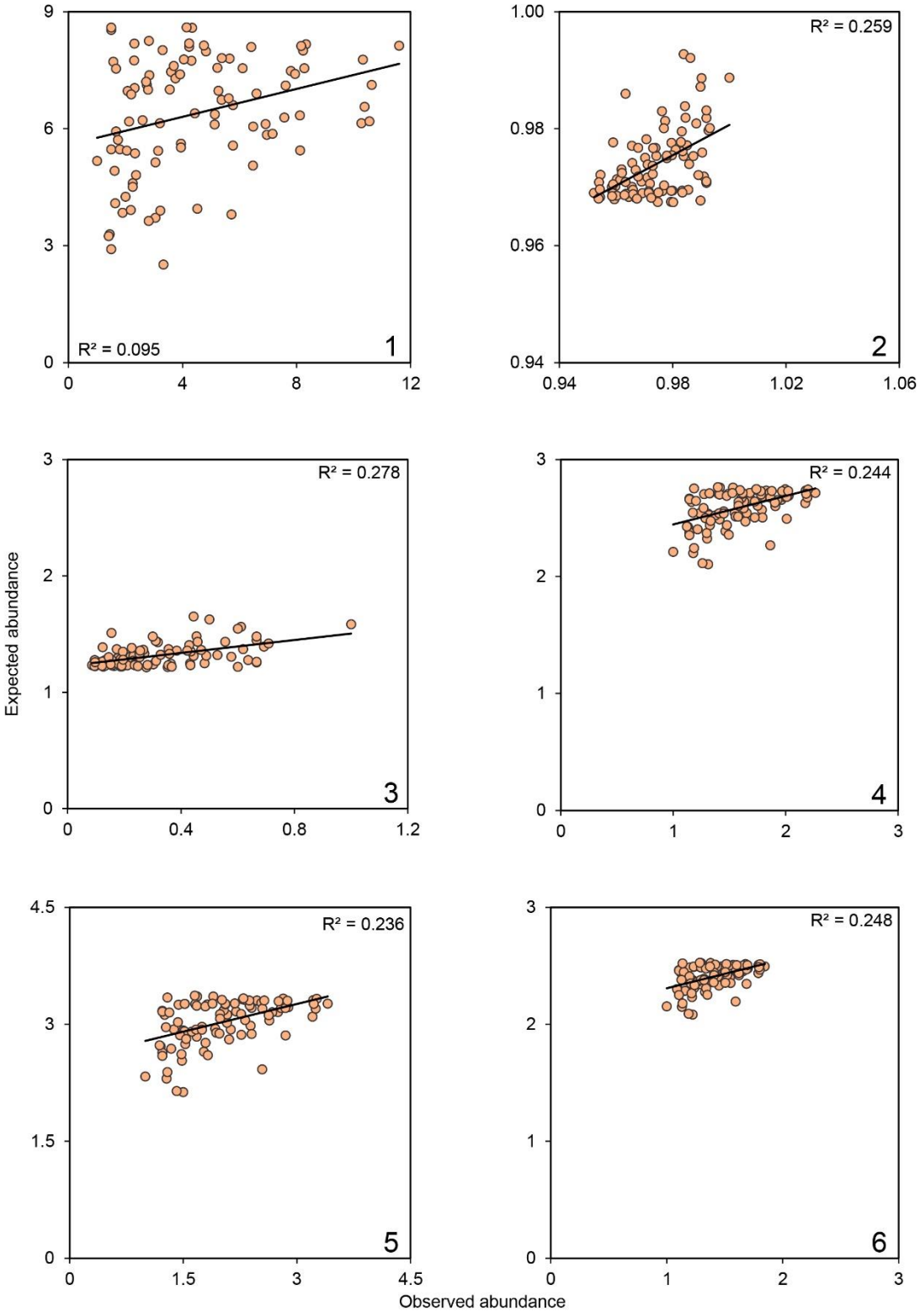
ANNEX 3 – Spatial representation of the first 10 principal components (PCs) estimated by PCA, including habitat (table 2), climate variables, hydrographic network and Terrain ruggedness index. Variable contribution for each PC are represented on table 4.



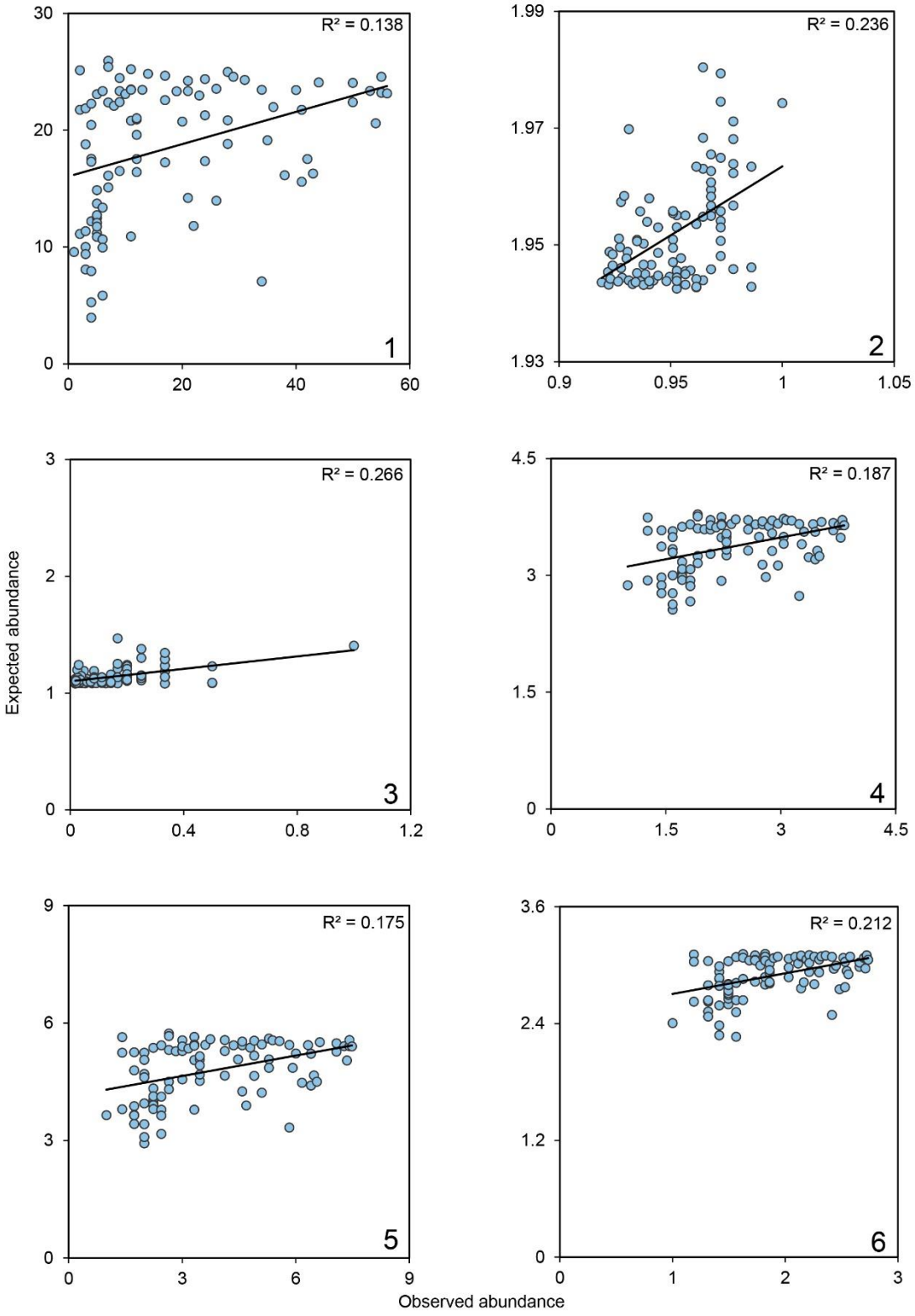
ANNEX 4 – Comparison between models performed with all the data, without transformations and with transformed data for the breeding season. 1 – original data; 2 – Box-Cox ($\lambda$) $= \frac{x^{\lambda}-1}{\lambda}$ ($\lambda$ = 0.101); 3 – reciprocal transformation ($1/x$); 4 – cube root ($\sqrt[3]{x}$); 5 – square root ($\sqrt{x}$); square root of the square root ($\sqrt{\sqrt{x}}$).



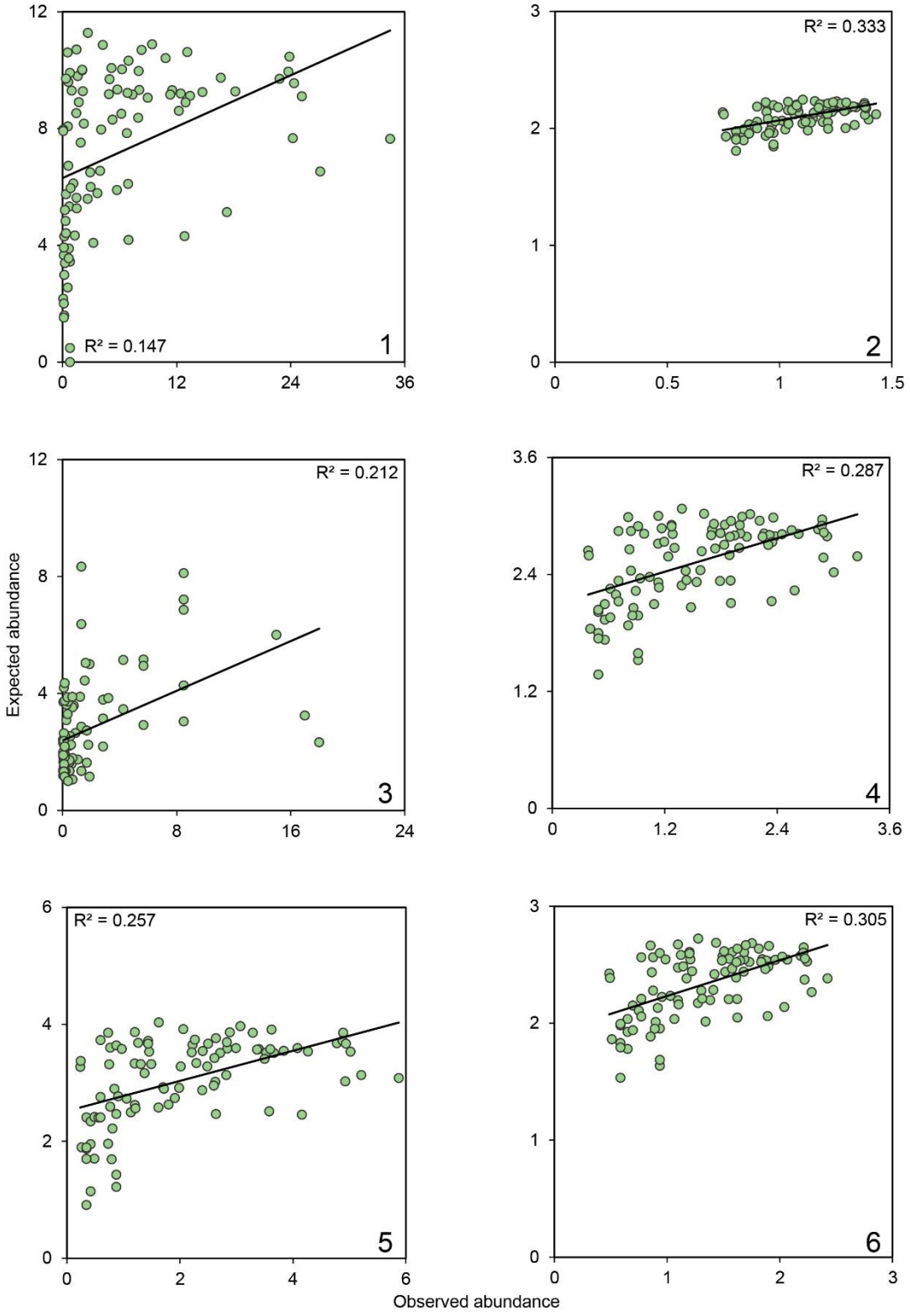
ANNEX 5 – Comparison between models performed the *Truncated data*, without transformations and with transformed data for the breeding season. 1 – original data; 2 – Box-Cox $(\boldsymbol{\lambda}) = \frac{x^{\lambda}-1}{\lambda}$ (λ = 0.101); 3 – reciprocal transformation $(1/x)$; 4 – cube root $(\sqrt[3]{x})$; 5 – square root $(\sqrt{x})$; square root of the square root $(\sqrt{\sqrt{x}})$.



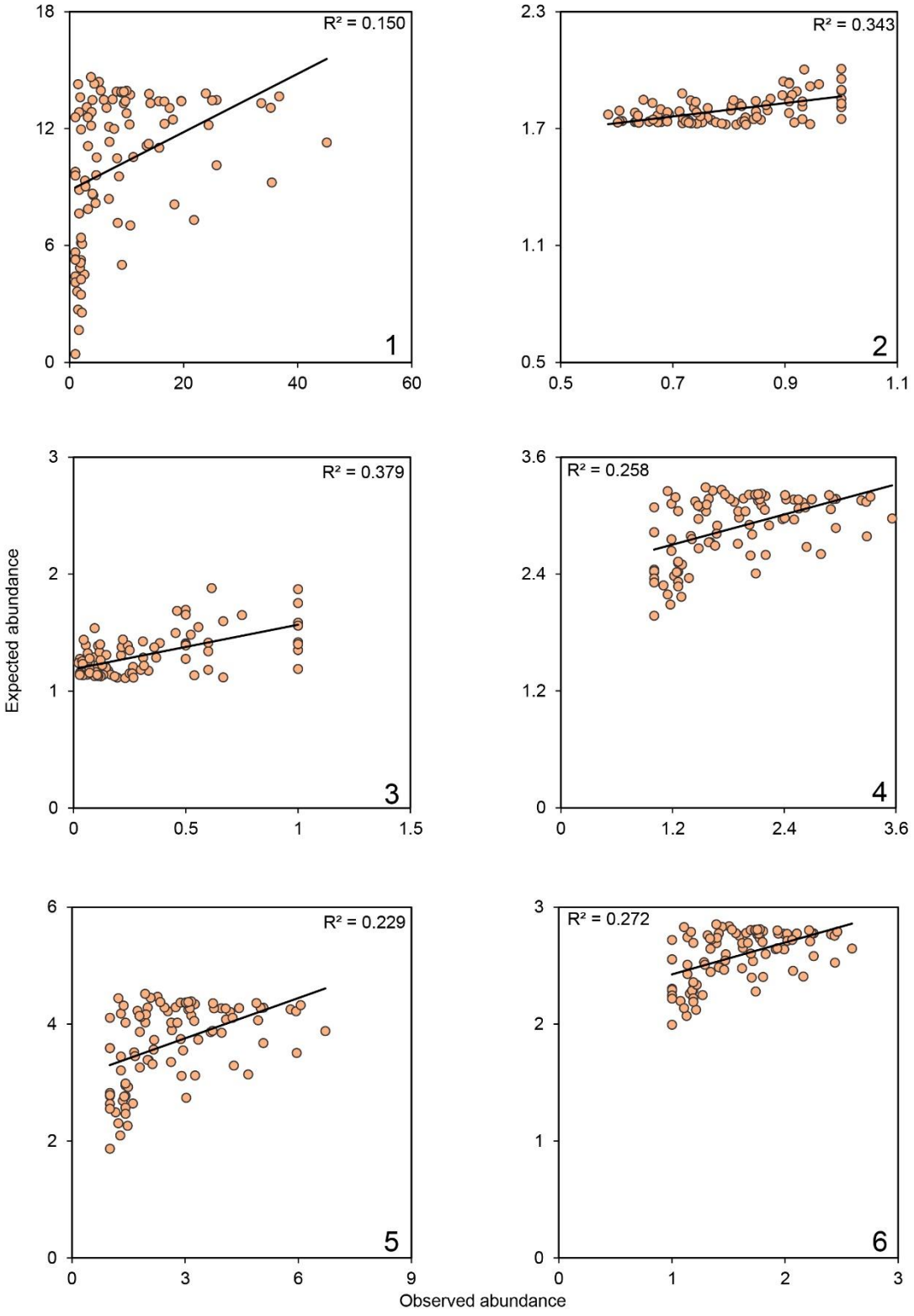
ANNEX 6 – Comparison between models performed with the maximum values, without transformations and with transformed data for the breeding season. 1 – original data; 2 – Box-Cox $x(\lambda) = \frac{x^{\lambda}-1}{\lambda}$; 3 – reciprocal transformation $(1/x)$; 4 – cube root $(\sqrt[3]{x})$; 5 – square root $(\sqrt{x})$; square root of the square root $(\sqrt{\sqrt{x}})$.



ANNEX 7 – Comparison between models performed with all the data, without transformations and with transformed data for the non-breeding season. 1 – original data; 2 – Box-Cox $x(\boldsymbol{\lambda}) = \frac{x^{\lambda}-1}{\boldsymbol{\lambda}}$; 3 – reciprocal transformation ($1/x$); 4 – cube root ($\sqrt[3]{x}$); 5 – square root ($\sqrt{x}$); square root of the square root ($\sqrt{\sqrt{x}}$).



ANNEX 8 – Comparison between models performed the *Truncated data*, without transformations and with transformed data for the non-breeding season. 1 – original data; 2 – Box-Cox $x(\lambda) = \frac{x^{\lambda}-1}{\lambda}$; 3 – reciprocal transformation $(1/x)$; 4 – cube root $(\sqrt[3]{x})$; 5 – square root $(\sqrt{x})$; square root of the square root $(\sqrt{\sqrt{x}})$.



ANNEX 9 – Comparison between models performed with the maximum values, without transformations and with transformed data for the non-breeding season. 1 – original data; 2 – Box-Cox $x(\lambda) = \frac{x^{\lambda}-1}{\lambda}$; 3 – reciprocal transformation $(1/x)$; 4 – cube root $(\sqrt[3]{x})$; 5 – square root $(\sqrt{x})$; square root of the square root $(\sqrt{\sqrt{x}})$.